BTP500
Data structures and Algorithms

WEEK- 1
C++ concepts

A Brief Introduction to C++

In this topic we will see:

- The similarities between C# and C++
- Some differences, including:
  - Global variables and functions
  - The preprocessor, compilation, namespaces
  - Printing
- Concluding with
  - Classes, templates
  - Pointers
  - Memory allocation and deallocation

Control Statements

All control statements are similar

```
if ( statement ) {
    // ...
} else if ( statement ) {
    // ...
} else {
    // ...
}
```

```
while ( statement ) {
    // ...
}
```

```
for ( int i = 0; i < N; ++i ) {
    // ...
}
```

```
do {
    // ...
} while ( statement );
```

Operators

Operators have similar functionality for built-in datatypes:

- Assignment `=`
- Arithmetic `+` `-` `*` `/` `%`
  `+=` `-=` `*=` `/=` `%=`
- Autoincrement `++`
- Autodecrement `--`
- Logical `&&` `||` `!`
- Relational `==` `!=` `<` `<=` `>=` `>`
- Comments `/*` `*/`
  `// to end of line`
- Bitwise `&` `|` `^` `~`
  `&=` `|=` `^=`
- Bit shifting `<<` `>>`
  `<<=` `>>=`

Arrays

Definition:

The *capacity* of an array is the entries it can hold

The *size* of an array is the number of useful entries

Accessing arrays is similar:

```
const int ARRAY_CAPACITY = 10; // prevents
 reassignment
int array[ARRAY_CAPACITY];


array[0] = 1;
for ( int i = 1; i < ARRAY_CAPACITY; ++i ) {
    array[i] = 2*array[i – 1] + 1;
}
```

Recall that arrays go from **`0`** to **`ARRAY_CAPACITY – 1`**

Functions

Function calls are similar, however, the are not required to be part of a class:

```
#include <iostream>
using namespace std;

// A function with a global name
int sqr( int n ) {
    return n*n;
}

int main() {
    cout << "The square of 3 is " << sqr(3) << endl;
    return 0;
}
```

C++/C# Differences

We will look at categories of differences between C++ and C#:

- Including header files (the preprocessor)
- The file is the base of compilation
- Namespaces
- Printing

The C++ Preprocessor

C++ is based on C, which was written in the early 1970s

Any command starting with a **#** in the first column is not a C/C++ statement, but rather a preprocessor statement

- The preprocessor performed very basic text-based (or *lexical*) substitutions
- The output is sent to the compiler

The C++ Preprocessor

The sequence is:

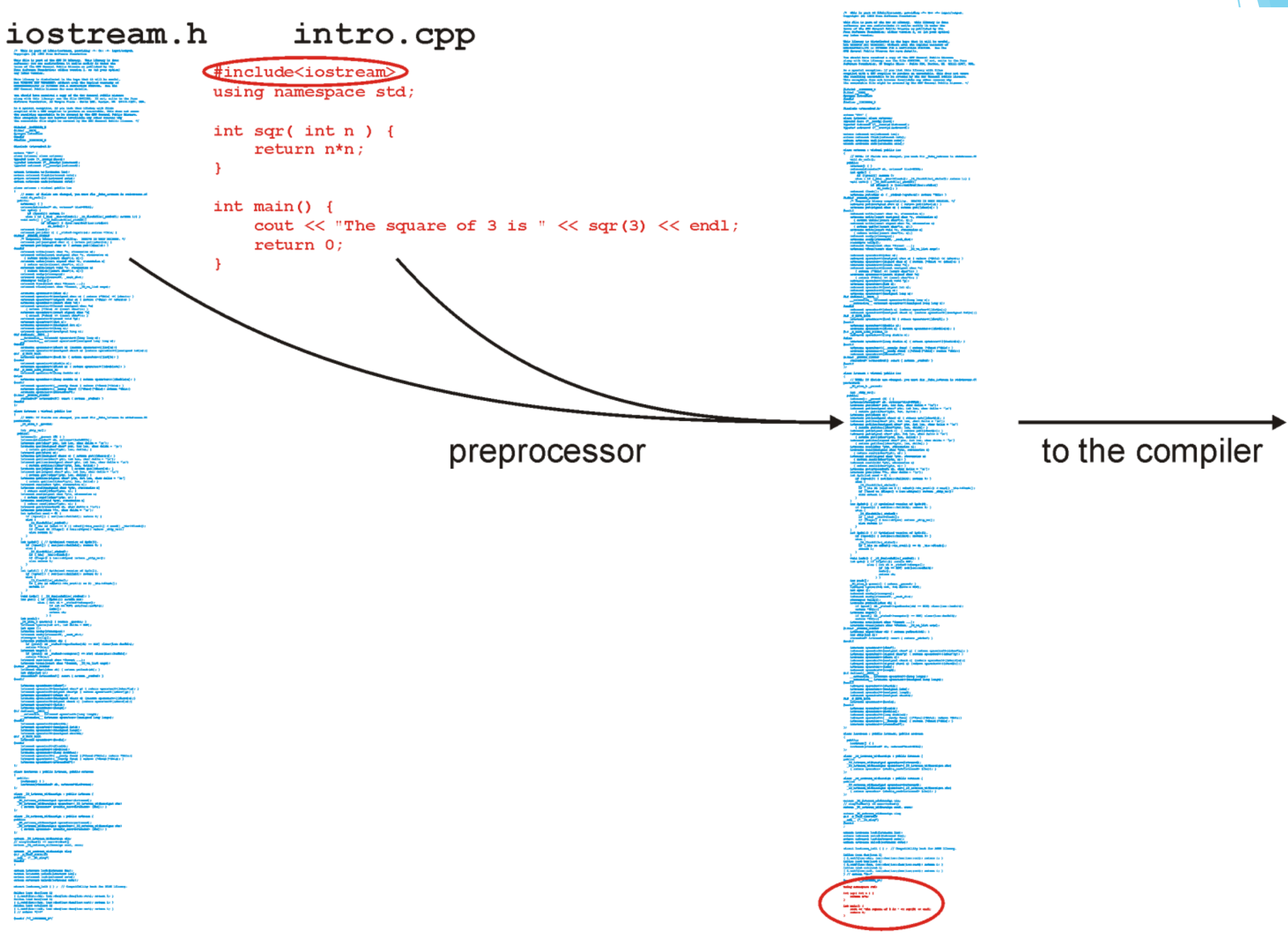
file (`filename.cpp`) → preprocessor → compiler (g++)

Note, this is done automatically by the compiler: no additional steps are necessary

At the top of any C++ program, you will see one or more directives starting with a **#**, *e.g.*,

```
#include <iostream>
```

The C++ Preprocessor



Libraries

You will note the difference:

```
#include <iostream>
#include "Single_list.h"
```

The first looks for a file `iostream.h` which is shipped with the compiler (the standard library)

The second looks in the current directory

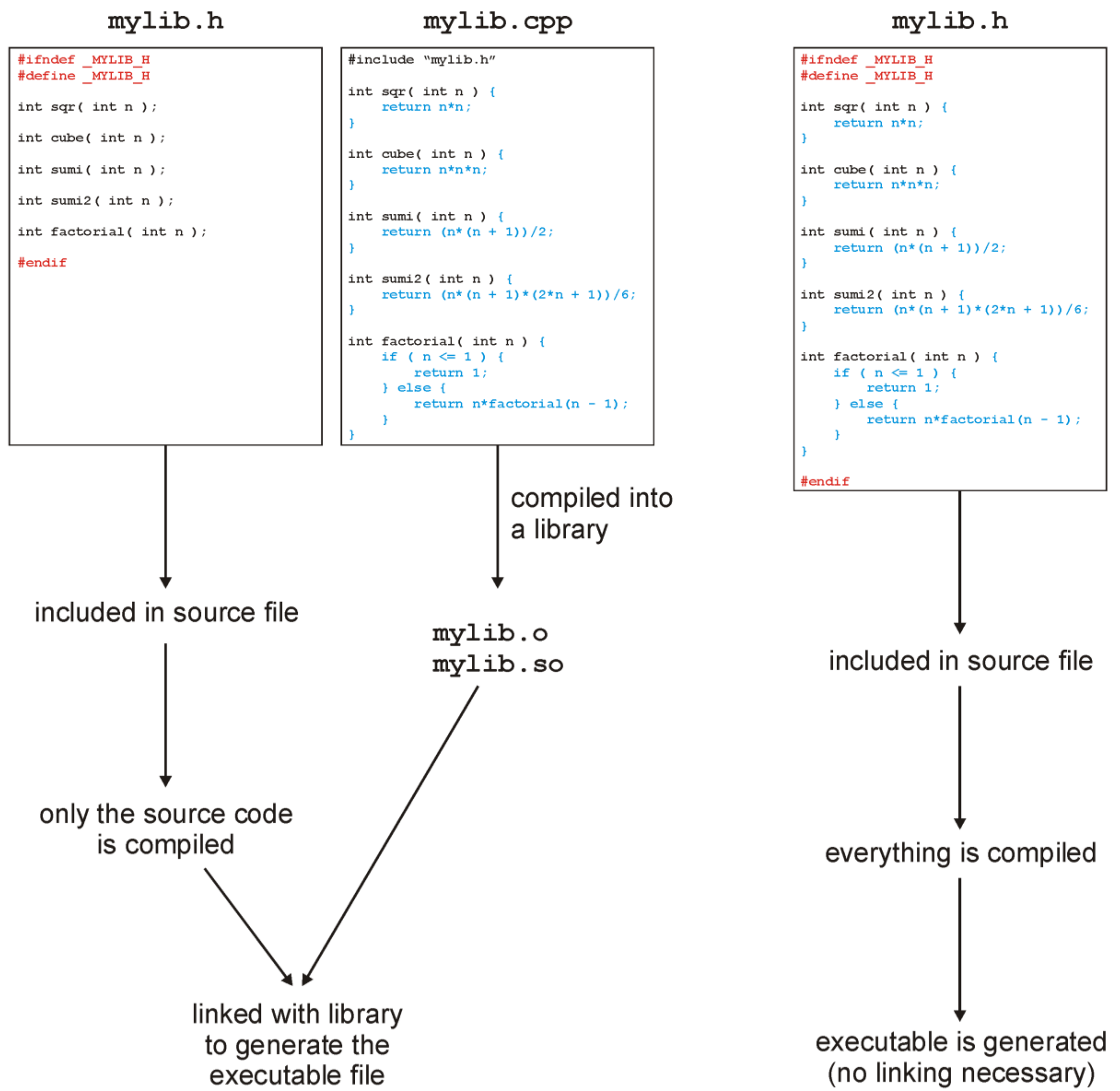
Libraries

In this class, you will put all code in the header file

This is not normal practice:

- Usually the header (`.h`) file only contains declarations
- The definitions (the actual implementations) are stored in a related file and compiled into an object file

The C++ Preprocessor



The C++ Preprocessor

With all these includes, it is always necessary to avoid the same file being included twice, otherwise you have duplicate definitions

This is done with guard statements:

```
#ifndef SINGLE_LIST_H
#define SINGLE_LIST_H

template <typename Type>
class Single_list {
  ///...
};

#endif
```

The C++ Preprocessor

This class definition contains only the signatures (or *prototypes*) of the operations

The actual member function definitions may be defined elsewhere, either in:

- The same file, or
- Another file which is compiled into an object file

We will use the first method

The File as the Unit of Compilation

Another difference is the unit of compilation

In C#, the class was the basis of compiling executable code:

```
class TestProgram {
    public static void Main() {
        System.Console.WriteLine( "Hello World" );
    }

}
```

The existence of a function with the signature

```
public static void Main();
```

determines whether or not a class can be compiled into an executable

The File as the Unit of Compilation

In C/C++, the file is the base unit of compilation:

- Any `.cpp` file may be compiled into object code
- Only files containing an `int main()` function can be compiled into an executable

The signature of main is:

```
int main () {
    // does some stuff
    return 0;
}
```

The operating system is expecting a return value

- Usually 0

The File as the Unit of Compilation

This file (**`example.cpp`**) contains two functions

```
#include<iostream>
using namespace std;

int sqr( int n ) {     // Function declaration and definition
    return n*n;
}

int main() {
    cout << "The square of 3 is " << sqr(3) << endl;
    return 0;
}
```

The File as the Unit of Compilation

This is an alternate form:

```
#include<iostream>
using namespace std;

int sqr( int );                // Function declaration

int main() {
    cout << "The square of 3 is " << sqr(3) << endl;
    return 0;
}

int sqr( int n ) {             // Function definition
    return n*n;                // The definition can be in another file
}
```

Namespaces

Variables defined:

- In functions are *local variables*
- In classes are *member variables*
- Elsewhere are *global variables*

Functions defined:

- In classes are *member functions*
- Elsewhere are *global functions*

In all these cases, the keyword **`static`** can modify the scope

Namespaces

Global variables/variables cause problems, especially in large projects

- Hundreds of employees
- Dozens of projects
- Everyone wanting a function `init()`

In C++ (and XML), this is solved using namespaces

Namespaces

You will only need this for the standard name space

- All variables and functions in the standard library are in the **`std`** namespace

```
#include <iostream>
std::cout << "Hello world!" << std::endl;
```

```
#include <iostream>
using namespace std;                // never used in
production code

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

Printing

Printing in C++ is done through overloading the **<<** operator:

```
cout << 3;
```

If the left-hand argument of **<<** is an object of type **`ostream`** (output *stream*) and the right-hand argument is a **`double`**, **`int`**, **`string`**, etc., an appropriate function which prints the object is called

Printing

The format is suggestive of what is happening:

- The objects are being *sent* to the **`cout`** (*console output*) object to be printed

```
 cout << "The square of 3 is " << sqr(3) <<
endl;
```

The objects being printed are:

- a **`string`**
- an **`int`**
- a platform-independent end-of-line identifier

Printing

How does

```
      cout << "The square of 3 is " << sqr(3) <<
endl;
```

work?

This is equivalent to

```
((cout << "The square of 3 is ") << sqr(3)) <<
                                   endl;
```

where **<<** is an operation (like **+**) which prints the object and returns the **cout** object

Printing

Visually:

```
((cout << "The square of 3 is ") << sqr(3)) << endl;
```

print `"The square of 3 is "` and return `cout`

```
(                cout                 << sqr(3)) << endl;
```

print the result of `sqr(3)` and return `cout`

```
cout                                   << endl;
```

print an end-of-line character (and return `cout`)

```
cout;
```

Printing

Another way to look at this is that

```
cout << "The square of 3 is " << sqr(3) << endl;
```

is the same as:

```
operator<<( operator<<( operator<<( cout, "The square of 3 is " ), sqr(3) ), endl );
```

This is how C++ treats these anyway...

Introduction to C++

The next five topics in C++ will be:

- Classes
- Templates
- Pointers
- Memory allocation
- Operator overloading

Classes

To begin, we will create a complex number class

To describe this class, we could use the following words:

- Store the real and imaginary components
- Allow the user to:
  - Create a complex number
  - Retrieve the real and imaginary parts
  - Find the absolute value and the exponential value
  - Normalize a non-zero complex number

Classes

An example of a C++ class declaration is:

```
class Complex {
    private:
        double re, im;

    public:
        Complex( double = 0.0, double = 0.0 );

        double real() const;
        double imag() const;
        double abs() const;
        Complex exp() const;

        void normalize();
};
```

Classes

This only declares the class structure

- It does not provide an implementation

We could, like C#, include the implementation in the class declaration, however, this is not, for numerous reasons, standard practice

The Complex Class

The next slide gives both the declaration of the `Complex` class as well as the associated definitions

- The assumption is that this is within a single file

The Complex Class

```
#ifndef _COMPLEX_H
#define _COMPLEX_H
#include <cmath>
class Complex {
    private:
        double re, im;

    public:
        Complex( double = 0.0, double = 0.0 );

        // Accessors
        double real() const;
        double imag() const;
        double abs() const;
        Complex exp() const;

        // Mutators
        void normalize();
};
```

The Complex Class

Associates functions back to the class

```
// Constructor
Complex::Complex( double r, double i ):
re( r ),
im( i ) {
    // empty constructor
}
```

Each member variable should be assigned

The order must be the same as the order in which the member variables are defined in the class

For built-in datatypes, this is a simple assignment. For member variables that are objects, this is a call to a constructor.

For built-in datatypes, the above is equivalent to:

```
// Constructor
Complex::Complex( double r, double i ):re( 0 ), im( 0 ) {
    re = r;
    im = i;
}
```

The Complex Class

```
// return the real component
double Complex::real() const {
    return re;
}

// return the imaginary component
double Complex::imag() const {
    return im;
}

// return the absolute value
double Complex::abs() const {
    return std::sqrt( re*re + im*im );
}
```

Refers to the member variables `re` and `im` of this class

The Complex Class

```
// Return the exponential of the complex value
Complex Complex::exp() const {
    double exp_re = std::exp( re );

    return Complex( exp_re*std::cos(im), exp_re*std::sin(im) );
}
```

The Complex Class

```
// Normalize the complex number (giving it unit absolute value, |z|
    = 1)
void Complex::normalize() {
    if ( re == 0 && im == 0 ) {
        return;
    }

    double absval = abs();
    re /= absval;
    im /= absval;
}

#endif
```

This calls the member function `double abs() const` from the `Complex` class on the object on which `void normalize()` was called

Visibility

Visibility in C# and Java is described by placing public/private/protected in front of each class member or member function

In C++, this is described by a block prefixed by one of

```
private:
protected:
public:
```

Visibility

```
class Complex {
    private:
        double re, im;

    public:
        Complex( double, double );

        double real() const;
        double imag() const;
        double abs() const;
        Complex exp() const;

        void normalize();
};
```

Visibility

The reason for the change in Java/C# was that the C++ version has been noted to be a source of errors

Code could be cut-and-paste from one location to another, and a poorly placed paste could change the visibility of some code:

`public` → `private` automatically caught

`private` → `public` difficult to catch and dangerous

Visibility

It is possible for a class to indicate that another class is allowed to access its **private** members

If class `ClassX` declares class `ClassY` to be a friend, then class `ClassY` can access (and modify) the private members of `ClassX`

Visibility

```
class ClassY;                   // declare that ClassY is a class

class ClassX {
    private:
        int privy;              // the variable privy is private

    friend class ClassY;        // ClassY is a "friend" of ClassX
};

class ClassY {                  // define ClassY
    private:
        ClassX value;           // Y stores one instance of X
    public:
        void set_x() {
            value.privy = 42;   // a member function of ClassY can
        }                       // access and modify the private
};                              // member privy of "value"
```

Pointers

One of the simplest ideas in C, but one which most students have a problem with is a pointer

- Every variable (barring optimization) is stored somewhere in memory
- That address is an integer, so why can't we store an address in a variable?

http://xkcd.com/138/

Pointers

We could simply have an ‘address’ type:

```
address ptr;     // store an address
                 // THIS IS WRONG
```

however, the compiler does not know what it is an address of (is it the address of an int, a double, etc.)

Instead, we have to indicate what it is pointing to:

```
int *ptr;    // a pointer to an integer
             // the address of the integer
variable 'ptr'
```

Pointers

First we must get the address of a variable

This is done with the **&** operator

(***a***mpersand/***a***ddress of)

For example,

```
int m = 5;     // m is an int storing 5
int *ptr;      // a pointer to an int
ptr = &m;      // assign to ptr the
               // address of m
```

Pointers

We can even print the addresses:

```
int m = 5;      // m is an int storing 5
int *ptr;       // a pointer to an int
ptr = &m;       // assign to ptr the
                // address of m
cout << ptr << endl;
```

prints `0xffffd352`, a 32-bit number

- The computer uses 32-bit addresses

Pointers

We have pointers: we would now like to manipulate what is stored at that address

We can access/modify what is stored at that memory location by using the * operator (dereference)

```
int m = 5;
int *ptr;
ptr = &m;
cout << *ptr << endl;  // prints 5
```

Pointers

Similarly, we can modify values stored at an address:

```
int m = 5;
int *ptr;
ptr = &m;

lo
```

http://xkcd.com/371/

Pointers

Pointers to objects must, similarly be dereferenced:

```
Complex z( 3, 4 );
Complex *pz;
pz = &z;
cout << z.abs() << endl;
cout << (*pz).abs() << endl;
```

Pointers

One short hand for this is to replace

```
(*pz).abs();
```

with

```
pz->abs();
```

Memory Allocation

Memory allocation in C++ is done through the `new` operator

This is an explicit request to the operating system for memory

- This is a very expensive operation
- The OS must:
  - Find the appropriate amount of memory,
  - Indicate that it has been allocated, and
  - Return the address of the first memory location

Memory Allocation

Memory deallocation differs, however:

- C# uses automatic garbage collection
- C++ requires the user to explicitly deallocate memory

Note however, that:

- *managed C*++ has garbage collection
- other tools are also available for C++ to perform automatic garbage collection

Memory Allocation

Inside a function, memory allocation of declared variables is dealt with by the compiler

```
int my_func() {
    Complex<double> z(3, 4);  // calls constructor with 3, 4
                              // creates 3 + 4j
                        // 16 bytes are allocated by the compiler

    double r = z.abs(); // 8 bytes are allocated by the compiler

    return 0;           // The compiler reclaims the 24 bytes
}
```

Memory Allocation

Memory for a single instance of a class (one object) is allocated using the new operator, *e.g.*,

```
Complex<double> *pz = new Complex<double>( 3,
4 );
```

The new operator returns the address of the first byte of the memory allocated

Memory Allocation

We can even print the address to the screen

If we were to execute

```
cout << "The address pz is " << pz << endl;
```

we would see output like:

```
The address pz is 0x00ef3b40
```

Memory Allocation

Next, to deallocate the memory (once we're finished with it) we must explicitly tell the operating system using the delete operator:

```
delete pz;
```

Memory Allocation

Consider a linked list where each node is allocated:

```
new Node<Type>( obj )
```

Such a call will be made each time a new element is added to the linked list

For each `new`, there must be a corresponding `delete`:

- Each removal of an object requires a call to `delete`
- If a non-empty list is itself being deleted, the destructor must call `delete` on all remaining nodes