# C/C++ PRIMER

# LECTURE 6: RESOURCE ACQUISITION (RAII) INHERITANCE AND OPERATOR OVERLOADING

Fabian Wermelinger

Harvard University

### **OUTLINE**

- Resource acquisition (RAII)
- Operator overloading
- Inheritance and the this pointer

Resource Acquisition Is Initialization (RAII), is a C++ programming technique which binds the *life cycle of a resource* that must be acquired before use (allocated heap memory, thread of execution, open socket, open file, locked mutex, disk space, database connection—anything that exists in limited supply) to the lifetime of an object (https://en.cppreference.com/w/cpp/language/raii).

- Encapsulate each resource into a class.
- The class constructor acquires the resource and constructs other attributes or throws an exception if that can not be done.
- The destructor releases the resource and never throws an exception.
- The resource is then used through a class instance with **bounded lifetime**. E.g. automatic variable inside a function body.

This ensures that when something goes wrong in a later part of the code which depends on the acquired resource (e.g. something else throws an exception but we continue running the code), the resource will still be freed when the end of lifetime has been reached.

#### **Examples:**

- Container classes in the Standard Template Library (STL)
- Anything that automatically allocates dynamic memory on the heap and automatically frees the resource again when the object reaches its end of life. This prevents memory leaks.
- In C, memory management is performed explicitly. It is sometimes not clear who is the *owner* of the resource and premature release of the resource leads to bugs in the code (often hard to find).

#### The rule of three:

- If your class follows the RAII principles (resource management) then the rule of three states that if you define one of the following, you should define all of them (the rule of three)
  - 1. Destructor
  - 2. Copy constructor
  - 3. Copy assignment operator
  - We have seen the first two last time and will talk about the third today.
- With the C++11 standard, move semantics have been introduced. Move semantics allow to transfer ownership of resources between instances of the class. They extend the rule of three to the rule of five.

#### The rule of five:

- With the C++11 standard, move semantics have been introduced. Move semantics allow to *transfer ownership* of resources between instances of the class. They extend the rule of three to the *rule of five*:
  - 1. Destructor
  - 2. Copy constructor
  - 3. Copy assignment operator
  - 4. Move constructor
  - 5. Move assignment operator

We will not talk about move semantics in this class.

In C++ you can do things like this:

```
1 class Foo
2 {
3     float data_[100];
4
5 public:
6     float &get(int i) { return data_[i]; }
7     float get(int i) const { return data_[i]; }
8 };
9
10 int main(void)
11 {
12     Foo f;
13     f.get(0) = 1;
14     return 0;
15 }
```

What do you think is happening in line 13? Which of the two methods in line 6 and line 7 is called? 

©

The method returns a reference as an lvalue for which it is legal to assign another value to it. We will not go into depth of value categories in C++, a good blog post to read is given at the bottom of this slide.

An Ivalue has *identity* and appears on the *left* side in an assignment. An rvalue has no identity (a temporary value) and appears on the right side of an assignment.

```
1 class Foo
2 {
3     float data_[100];
4
5 public:
6    float &get(int i) { return data_[i]; }
7     float get(int i) const { return data_[i]; }
8 };
```

Is the return value of the highlighted method an lvalue or an rvalue?

https://blog.knatten.org/2018/03/09/Ivalues-rvalues-glvalues-prvalues-xvalues-help https://en.cppreference.com/w/cpp/language/value\_category

Returning references from member functions has two use cases:

- 1. Change class state through lvalue assignment (what we just discussed).
- 2. Return a reference of something heavy that does not need to be copied. The following code illustrates this:

Note how the const type qualifier is used, what does it express?

References as return arguments in methods allow for an elegant way to overload operators.

#### Example: indexing operator[]

```
1 #include <cassert>
   class Foo
       float data_[100];
   public:
        float &operator[](unsigned int i) // returns lvalue
            assert(i < 100); // using assertions is good practice!</pre>
10
            return data_[i];
11
12
        }
13
       float operator[](unsigned int i) const // returns rvalue
14
15
            assert(i < 100); // using assertions is good practice!</pre>
16
            return data_[i];
17
18
19 };
20
21 int main(void)
22 {
23
       Foo f;
24
       f[0] = 1;  // calls float &operator[](unsigned int i)
       float a = f[0]; // calls float operator[](unsigned int i) const
25
26
       return 0;
27 }
```

**Example: copy assignment operator=** (this one is tricky)

Recall: the rule of three: destructor, copy constructor and copy assignment operator (operator = )

```
class Foo
        float data_[10];
   public:
        Foo & operator = (const Foo & other)
            if (&other == this) {
 8
                return *this;
10
            for (int i = 0; i < 10; ++i) {
11
12
                data_[i] = other.data_[i];
13
14
            return *this;
15
16 };
```

- The other parameter is passed as a const reference.
- You must check for self assignment. (More recent C++ idioms may use a swap operation instead. We do not discuss that here.)
- You must return a reference to the calling instance. (Why?)

**Example: copy assignment operator=** (this one is tricky)

```
1 class Foo
       float data_[10];
 4
   public:
       Foo(const int v) { /* pass */ }
       Foo &operator=(const Foo &c)
           if (&c == this) {
10
11
                return *this;
12
           for (int i = 0; i < 10; ++i) {
13
                data_[i] = c.data_[i];
14
15
           return *this;
17
18 };
```

What do you think is being called in the following code?

```
1 int main(void)
2 {
3    Foo f = 1; // what is being called here?
4    Foo g(2); // what is being called here?
5    f = g; // what is being called here?
6    return 0;
7 }
```

#### **Example: copy assignment operator=** (this one is tricky)

- If you do not provide a destructor, copy constructor or copy assignment operator in your class, then the compiler will generate a default version for you.
- Most of the time the default version does not what you need it to do!
- Example: Dynamic memory on the heap. The default copy constructor and copy assignment operator will copy the value of the pointer and not the underlying data (shallow copy). In that case the copied object shares the data with the owner object. In most cases this will be a bug!
- For RAII, you *must* provide implementations for a destructor, copy constructor and copy assignment operator (and move constructor and move assignment operator in case you need move semantics).

See also: https://en.cppreference.com/w/cpp/language/operators

Hands-On: Copy assignment operator (hands-on/01/README.md)

The code in copy\_assignment\_bug.cpp generates the following output

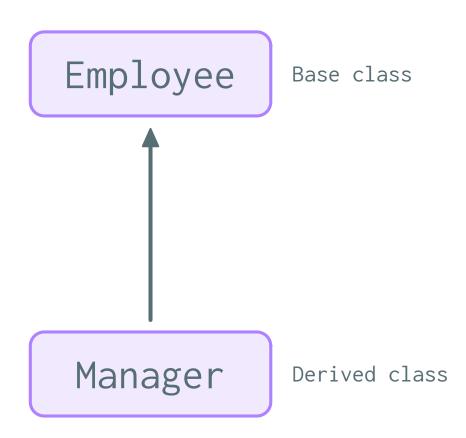
while we would expect the following output

We will work on this problem together and we need a debugger to get a deeper insight. You can install for example gdb

(https://www.gnu.org/software/gdb/) via sudo apt-get install gdb.

- One of the key elements in OOP is inheritance: a new class can inherit the data members and member functions of another already defined class.
- The derived class may add new data members or methods, shadow members or extend functionality through polymorphism.
- Examples:
  - Student, Assistant and Professor are University members
  - Triangles, Circles, Rectangles are shapes
  - Dogs, Cats, Mice and Elephants are animals
- Inheritance allows for a modular design by encapsulating data and sharing functionality that can further be extended or modify where needed. The extension of a feature *does not* duplicate code in the design hierarchy.

- We work through an example with base class Employee and derived class Manager.
- Keep in mind that every Manager is an Employee but not every Employee is a Manager.
- A derived object can be treated the same as an object of its base class. We see what this means when we talk about base class pointers.
- A derived class can not access the private member of its base class.
- A derived class constructor has to call one of the constructors for its base class. Default constructor is called otherwise.



#### The base class: Employee

```
1 #include <iostream>
2 #include <string>
3
4 class Employee
5 {
6 public:
7    Employee(std::string name) : name_(name) {}
8
9    void print_info() const { std::cout << "Name: " << name_ << std::endl; }
10    std::string get_name() const { return name_; }
11
12 private:
13    std::string name_;
14 };</pre>
```

#### Recall:

- private: can not be accessed from outside the class.
- public: can be accessed from outside the class.

The derived class: Manager

```
Employee(std::string name) : name_(name) {}
15 class Manager : public Employee
16 {
17 public:
       Manager(std::string name, std::string department)
            : Employee(name), department_(department)
19
        {
}
21
22
23
       void print_info() const
24
25
            Employee::print_info();
            std::cout << "Head of the " << department_ << std::endl;</pre>
27
29 private:
30
        std::string department_;
31 };
```

#### The derived class: Manager

```
Employee(std::string name) : name_(name) {}
           std::cout << "Name: " << name_ << std::endl;</pre>
  class Manager : public Employee
       Manager(std::string name, std::string department)
18
19
           : Employee(name), department_(department)
           Employee::print info();
           std::cout << "Head of the " << department << std::endl;
  private:
       std::string department_;
```

- Manager inherits public access rights. Recall that public and private allow or disallow access to the class members. The protected access modifier allows derived classes access to members that are protected from the public (similar to private but less strict).
- If you call a constructor for the base class, you must call it in the initializer list.
- Additional data members can be defined in derived classes with appropriate access modifiers.

#### The derived class: Manager

```
void print_info() const {
            std::cout << "Name: " << name_ << std::endl;</pre>
   class Manager : public Employee
       Manager(std::string name, std::string department)
            : Employee(name), department_(department)
23
       void print_info() const
24
25
            Employee::print_info();
            std::cout << "Head of the " << department_ << std::endl;</pre>
27
```

- We can overload member functions in the derived class.
   (Note the emphasis on overload.)
- The derived class calls the print\_info() explicitly via the scope operator ":: ", that is Employee::print\_info().
- After that call it adds one more line of output. Manager extends the functionality of print\_info() in the Employee class.

The derived class: Manager

```
int main(void)

{
    Employee e("Simpson");
    Manager m("Mr. Burns", "Nuclear Power Plant");
    e.print_info();
    m.print_info();
    return 0;
}
```

What output do you expect?

#### Access modifiers revisited:

- The usage of public, protected and private within the class definition should now be clear.
- Naturally they are linked to inheritance:
  - The access modifiers keep their rights as they were.

```
1 class Derived : public Base
```

The inherited members with public access become protected in the derived class.

```
1 class Derived : protected Base
```

The inherited members with public and protected access become private in the derived class.

```
1 class Derived : private Base
```

#### Access modifiers revisited: examples

Base class Public inheritance Protected inheritance Private inheritance

```
1 class Base
                               1 class Deriv1 : public Base 1 class Deriv2 :
                                                                                            1 class Deriv3 :
 2 {
                               2 {
                                                                    protected Base
                                                                                                   private Base
   public:
                                      void do_things()
                                                                                            3 {
       void pub() {}
                                                                    void do_things()
                                                                                                   void do_things()
   protected:
                                          pub(); // OK
       void pro() {}
                                          pro(); // OK
                                                                        pub(); // OK
                                                                                                       pub(); // OK
                                          pri(); // NOT OK
                                                                        pro(); // OK
                                                                                                       pro(); // OK
   private:
       void pri() {}
                                                                        pri(); // NOT OK
                                                                                                       pri(); // NOT OK
 9 };
                               9 };
                                                             10 };
                                                                                           10 };
11 int main(void)
                              11 int main(void)
                                                             11
                                                                                            11
                                                             12
12 {
                              12 {
                                                                                           12 class Deriv4:
       Base base;
                                      Deriv1 dpub;
                                                             13 int main(void)
                                                                                                   public Deriv3
13
                              13
                                                                                            13
       base.pub(); // OK
                                      dpub.pro(); // NOT OK 14 {
                                                                                           14 {
14
                              14
       base.pro(); // NOT OK 15
                                                                    Deriv2 dpro;
                                                                                                   void do_things()
15
                                                             15
                                                                                            15
       base.pri(); // NOT OK
                                      return 0;
                                                                    dpro.pub(); // NOT OK
16
                              16
                                                             16
17
                              17 }
                                                             17
                                                                                            17
                                                                                                       pub(); // NOT OK
       return 0;
                                                                    return 0;
                                                                                                       pro(); // NOT OK
                                                                                           18
                                                                                                       pri(); // NOT OK
19 }
                                                             19 }
                                                                                            20
                                                                                           21 };
```

#### The this pointer

- We have already seen the this pointer in an earlier lecture.
- It showed up again today in operator overloading.
- The value of the this pointer is the address of the *implicit object* parameter (the object on which non-static member functions are being called).
- It can appear in the following contexts:
  - 1. Within the body of non-static member functions.
  - 2. Within the declaration of non-static member functions after the type qualifier sequence.
  - 3. Within default member initialization.
- It is similar to the self reference in python except that this is implicit

#### The this pointer: examples

```
1 class Foo
 3
       int x;
       void bar()
 6
           x = 6; // implicit use of this-> (same as this->x = 6;)
           this->x = 5; // explicit use of this->
 9
10
11
       void bar() const // cv-qualifier
12
13
           x = 7; // NOT OK: *this is constant
14
15
16
       void bar(int x) // parameter x shadows the member with same name
17
            this->x = x; // this->x and x are not the same in this context!
18
19
20
21
       int y;
22
       Foo(int x)
23
            : x(x),
             v(this->x) // y is initialized using member x
24
25
26
27 };
```