

15.3 — Move constructors and move assignment

BY ALEX ON FEBRUARY 26TH, 2017 | LAST MODIFIED BY ALEX ON JANUARY 23RD, 2020

In lesson **15.1 -- Intro to smart pointers and move semantics**, we took a look at `std::auto_ptr`, discussed the desire for move semantics, and took a look at some of the downsides that occur when functions designed for copy semantics (copy constructors and copy assignment operators) are redefined to implement move semantics.

In this lesson, we'll take a deeper look at how C++11 resolves these problems via move constructors and move assignment.

Copy constructors and copy assignment

First, let's take a moment to recap copy semantics.

Copy constructors are used to initialize a class by making a copy of an object of the same class. Copy assignment is used to copy one class to another existing class. By default, C++ will provide a copy constructor and copy assignment operator if one is not explicitly provided. These compiler-provided functions do shallow copies, which may cause problems for classes that allocate dynamic memory. So classes that deal with dynamic memory should override these functions to do deep copies.

Returning back to our `Auto_ptr` smart pointer class example from the first lesson in this chapter, let's look at a version that implements a copy constructor and copy assignment operator that do deep copies, and a sample program that exercises them:

```

1  template<class T>
2  class Auto_ptr3
3  {
4      T* m_ptr;
5  public:
6      Auto_ptr3(T* ptr = nullptr)
7          :m_ptr(ptr)
8      {
9      }
10
11     ~Auto_ptr3()
12     {
13         delete m_ptr;
14     }
15
16     // Copy constructor
17     // Do deep copy of a.m_ptr to m_ptr
18     Auto_ptr3(const Auto_ptr3& a)
19     {
20         m_ptr = new T;
21         *m_ptr = *a.m_ptr;
22     }
23
24     // Copy assignment
25     // Do deep copy of a.m_ptr to m_ptr
26     Auto_ptr3& operator=(const Auto_ptr3& a)
27     {
28         // Self-assignment detection
29         if (&a == this)
30             return *this;
31
32         // Release any resource we're holding
33         delete m_ptr;
34

```

```

35     // Copy the resource
36     m_ptr = new T;
37     *m_ptr = *a.m_ptr;
38
39     return *this;
40 }
41
42 T& operator*() const { return *m_ptr; }
43 T* operator->() const { return m_ptr; }
44 bool isNull() const { return m_ptr == nullptr; }
45 };
46
47 class Resource
48 {
49 public:
50     Resource() { std::cout << "Resource acquired\n"; }
51     ~Resource() { std::cout << "Resource destroyed\n"; }
52 };
53
54 Auto_ptr3<Resource> generateResource()
55 {
56     Auto_ptr3<Resource> res(new Resource);
57     return res; // this return value will invoke the copy constructor
58 }
59
60 int main()
61 {
62     Auto_ptr3<Resource> mainres;
63     mainres = generateResource(); // this assignment will invoke the copy assignment
64
65     return 0;
66 }

```

In this program, we're using a function named `generateResource()` to create a smart pointer encapsulated resource, which is then passed back to function `main()`. Function `main()` then assigns that to an existing `Auto_ptr3` object.

When this program is run, it prints:

```

Resource acquired
Resource acquired
Resource destroyed
Resource acquired
Resource destroyed
Resource destroyed

```

(Note: You may only get 4 outputs if your compiler elides the return value from function `generateResource()`)

That's a lot of resource creation and destruction going on for such a simple program! What's going on here?

Let's take a closer look. There are 6 key steps that happen in this program (one for each printed message):

- 1) Inside `generateResource()`, local variable `res` is created and initialized with a dynamically allocated `Resource`, which causes the first "Resource acquired".
- 2) `res` is returned back to `main()` by value. We return by value here because `res` is a local variable -- it can't be returned by address or reference because `res` will be destroyed when `generateResource()` ends. So `res` is copy constructed into a temporary object. Since our copy constructor does a deep copy, a new `Resource` is allocated here, which causes the second "Resource acquired".
- 3) `res` goes out of scope, destroying the originally created `Resource`, which causes the first "Resource destroyed".

- 4) The temporary object is assigned to `mainres` by copy assignment. Since our copy assignment also does a deep copy, a new `Resource` is allocated, causing yet another “Resource acquired”.
- 5) The assignment expression ends, and the temporary object goes out of expression scope and is destroyed, causing a “Resource destroyed”.
- 6) At the end of `main()`, `mainres` goes out of scope, and our final “Resource destroyed” is displayed.

So, in short, because we call the copy constructor once to copy construct `res` to a temporary, and copy assignment once to copy the temporary into `mainres`, we end up allocating and destroying 3 separate objects in total.

Inefficient, but at least it doesn't crash!

However, with move semantics, we can do better.

Move constructors and move assignment

C++11 defines two new functions in service of move semantics: a move constructor, and a move assignment operator. Whereas the goal of the copy constructor and copy assignment is to make a copy of one object to another, the goal of the move constructor and move assignment is to move ownership of the resources from one object to another (which is much less expensive than making a copy).

Defining a move constructor and move assignment work analogously to their copy counterparts. However, whereas the copy flavors of these functions take a `const l-value` reference parameter, the move flavors of these functions use `non-const r-value` reference parameters.

Here's the same `Auto_ptr3` class as above, with a move constructor and move assignment operator added. We've left in the deep-copying copy constructor and copy assignment operator for comparison purposes.

```

1  #include <iostream>
2
3  template<class T>
4  class Auto_ptr4
5  {
6      T* m_ptr;
7  public:
8      Auto_ptr4(T* ptr = nullptr)
9          :m_ptr(ptr)
10     {
11     }
12
13     ~Auto_ptr4()
14     {
15         delete m_ptr;
16     }
17
18     // Copy constructor
19     // Do deep copy of a.m_ptr to m_ptr
20     Auto_ptr4(const Auto_ptr4& a)
21     {
22         m_ptr = new T;
23         *m_ptr = *a.m_ptr;
24     }
25
26     // Move constructor
27     // Transfer ownership of a.m_ptr to m_ptr
28     Auto_ptr4(Auto_ptr4&& a)
29         : m_ptr(a.m_ptr)
30     {
31         a.m_ptr = nullptr; // we'll talk more about this line below
32     }
33
34     // Copy assignment
```

```
35 // Do deep copy of a.m_ptr to m_ptr
36 Auto_ptr4& operator=(const Auto_ptr4& a)
37 {
38     // Self-assignment detection
39     if (&a == this)
40         return *this;
41
42     // Release any resource we're holding
43     delete m_ptr;
44
45     // Copy the resource
46     m_ptr = new T;
47     *m_ptr = *a.m_ptr;
48
49     return *this;
50 }
51
52 // Move assignment
53 // Transfer ownership of a.m_ptr to m_ptr
54 Auto_ptr4& operator=(Auto_ptr4&& a)
55 {
56     // Self-assignment detection
57     if (&a == this)
58         return *this;
59
60     // Release any resource we're holding
61     delete m_ptr;
62
63     // Transfer ownership of a.m_ptr to m_ptr
64     m_ptr = a.m_ptr;
65     a.m_ptr = nullptr; // we'll talk more about this line below
66
67     return *this;
68 }
69
70 T& operator*() const { return *m_ptr; }
71 T* operator->() const { return m_ptr; }
72 bool isNull() const { return m_ptr == nullptr; }
73 };
74
75 class Resource
76 {
77 public:
78     Resource() { std::cout << "Resource acquired\n"; }
79     ~Resource() { std::cout << "Resource destroyed\n"; }
80 };
81
82 Auto_ptr4<Resource> generateResource()
83 {
84     Auto_ptr4<Resource> res(new Resource);
85     return res; // this return value will invoke the move constructor
86 }
87
88 int main()
89 {
90     Auto_ptr4<Resource> mainres;
91     mainres = generateResource(); // this assignment will invoke the move assignment
92
93     return 0;
94 }
```

The move constructor and move assignment operator are simple. Instead of deep copying the source object (a) into the implicit object, we simply move (steal) the source object's resources. This involves shallow copying the source pointer into the implicit object, then setting the source pointer to null.

When run, this program prints:

```
Resource acquired  
Resource destroyed
```

That's much better!

The flow of the program is exactly the same as before. However, instead of calling the copy constructor and copy assignment operators, this program calls the move constructor and move assignment operators. Looking a little more deeply:

- 1) Inside `generateResource()`, local variable `res` is created and initialized with a dynamically allocated `Resource`, which causes the first "Resource acquired".
- 2) `Res` is returned back to `main()` by value. `Res` is move constructed into a temporary object, transferring the dynamically created object stored in `res` to the temporary object. We'll talk about why this happens below.
- 3) `Res` goes out of scope. Because `res` no longer manages a pointer (it was moved to the temporary), nothing interesting happens here.
- 4) The temporary object is move assigned to `mainres`. This transfers the dynamically created object stored in the temporary to `mainres`.
- 5) The assignment expression ends, and the temporary object goes out of expression scope and is destroyed. However, because the temporary no longer manages a pointer (it was moved to `mainres`), nothing interesting happens here either.
- 6) At the end of `main()`, `mainres` goes out of scope, and our final "Resource destroyed" is displayed.

So instead of copying our `Resource` twice (once for the copy constructor and once for the copy assignment), we transfer it twice. This is more efficient, as `Resource` is only constructed and destroyed once instead of three times.

When are the move constructor and move assignment called?

The move constructor and move assignment are called when those functions have been defined, and the argument for construction or assignment is an r-value. Most typically, this r-value will be a literal or temporary value.

In most cases, a move constructor and move assignment operator will not be provided by default, unless the class does not have any defined copy constructors, copy assignment, move assignment, or destructors. However, the default move constructor and move assignment do the same thing as the default copy constructor and copy assignment (make copies, not do moves).

Rule: If you want a move constructor and move assignment that do moves, you'll need to write them yourself.

The key insight behind move semantics

You now have enough context to understand the key insight behind move semantics.

If we construct an object or do an assignment where the argument is an l-value, the only thing we can reasonably do is copy the l-value. We can't assume it's safe to alter the l-value, because it may be used again later in the program. If we have an expression "`a = b`", we wouldn't reasonably expect `b` to be changed in any way.

However, if we construct an object or do an assignment where the argument is an r-value, then we know that r-value is just a temporary object of some kind. Instead of copying it (which can be expensive), we can simply transfer its resources (which is cheap) to the object we're constructing or assigning. This is safe to do because the temporary will be destroyed at the end of the expression anyway, so we know it will never be used again!

C++11, through r-value references, gives us the ability to provide different behaviors when the argument is an r-value vs an l-value, enabling us to make smarter and more efficient decisions about how our objects should behave.

Move functions should always leave both objects in a well-defined state

In the above examples, both the move constructor and move assignment functions set `a.m_ptr` to `nullptr`. This may seem extraneous -- after all, if "a" is a temporary r-value, why bother doing "cleanup" if parameter "a" is going to be destroyed anyway?

The answer is simple: When "a" goes out of scope, a's destructor will be called, and `a.m_ptr` will be deleted. If at that point, `a.m_ptr` is still pointing to the same object as `m_ptr`, then `m_ptr` will be left as a dangling pointer. When the object containing `m_ptr` eventually gets used (or destroyed), we'll get undefined behavior.

Additionally, in the next lesson we'll see cases where "a" can be an l-value. In such a case, "a" wouldn't be destroyed immediately, and could be queried further before its lifetime ends.

Automatic l-values returned by value may be moved instead of copied

In the `generateResource()` function of the `Auto_ptr4` example above, when variable `res` is returned by value, it is moved instead of copied, even though `res` is an l-value. The C++ specification has a special rule that says automatic objects returned from a function by value can be moved even if they are l-values. This makes sense, since `res` was going to be destroyed at the end of the function anyway! We might as well steal its resources instead of making an expensive and unnecessary copy.

Although the compiler can move l-value return values, in some cases it may be able to do even better by simply eliding the copy altogether (which avoids the need to make a copy or do a move at all). In such a case, neither the copy constructor nor move constructor would be called.

Disabling copying

In the `Auto_ptr4` class above, we left in the copy constructor and assignment operator for comparison purposes. But in move-enabled classes, it is sometimes desirable to delete the copy constructor and copy assignment functions to ensure copies aren't made. In the case of our `Auto_ptr` class, we don't want to copy our templated object `T` -- both because it's expensive, and whatever class `T` is may not even support copying!

Here's a version of `Auto_ptr` that supports move semantics but not copy semantics:

```

1  #include <iostream>
2
3  template<class T>
4  class Auto_ptr5
5  {
6      T* m_ptr;
7  public:
8      Auto_ptr5(T* ptr = nullptr)
9          :m_ptr(ptr)
10     {
11     }
12
13     ~Auto_ptr5()
14     {
15         delete m_ptr;
16     }
17
18     // Copy constructor -- no copying allowed!
19     Auto_ptr5(const Auto_ptr5& a) = delete;
20
21     // Move constructor
22     // Transfer ownership of a.m_ptr to m_ptr
23     Auto_ptr5(Auto_ptr5&& a)
```

```

24         : m_ptr(a.m_ptr)
25     {
26         a.m_ptr = nullptr;
27     }
28
29     // Copy assignment -- no copying allowed!
30     Auto_ptr5& operator=(const Auto_ptr5& a) = delete;
31
32     // Move assignment
33     // Transfer ownership of a.m_ptr to m_ptr
34     Auto_ptr5& operator=(Auto_ptr5&& a)
35     {
36         // Self-assignment detection
37         if (&a == this)
38             return *this;
39
40         // Release any resource we're holding
41         delete m_ptr;
42
43         // Transfer ownership of a.m_ptr to m_ptr
44         m_ptr = a.m_ptr;
45         a.m_ptr = nullptr;
46
47         return *this;
48     }
49
50     T& operator*() const { return *m_ptr; }
51     T* operator->() const { return m_ptr; }
52     bool isNull() const { return m_ptr == nullptr; }
53 };

```

If you were to try to pass an `Auto_ptr5` l-value to a function by value, the compiler would complain that the copy constructor required to initialize the copy constructor argument has been deleted. This is good, because we should probably be passing `Auto_ptr5` by const l-value reference anyway!

`Auto_ptr5` is (finally) a good smart pointer class. And, in fact the standard library contains a class very much like this one (that you should use instead), named `std::unique_ptr`. We'll talk more about `std::unique_ptr` later in this chapter.

Another example

Let's take a look at another class that uses dynamic memory: a simple dynamic templated array. This class contains a deep-copying copy constructor and copy assignment operator.

```

1  #include <iostream>
2
3  template <class T>
4  class DynamicArray
5  {
6  private:
7      T* m_array;
8      int m_length;
9
10 public:
11     DynamicArray(int length)
12         : m_array(new T[length]), m_length(length)
13     {
14     }
15
16     ~DynamicArray()
17     {
18         delete[] m_array;

```

```

19     }
20
21     // Copy constructor
22     DynamicArray(const DynamicArray &arr)
23         : m_length(arr.m_length)
24     {
25         m_array = new T[m_length];
26         for (int i = 0; i < m_length; ++i)
27             m_array[i] = arr.m_array[i];
28     }
29
30     // Copy assignment
31     DynamicArray& operator=(const DynamicArray &arr)
32     {
33         if (&arr == this)
34             return *this;
35
36         delete[] m_array;
37
38         m_length = arr.m_length;
39         m_array = new T[m_length];
40
41         for (int i = 0; i < m_length; ++i)
42             m_array[i] = arr.m_array[i];
43
44         return *this;
45     }
46
47     int getlength() const { return m_length; }
48     T& operator[](int index) { return m_array[index]; }
49     const T& operator[](int index) const { return m_array[index]; }
50
51 };

```

Now let's use this class in a program. To show you how this class performs when we allocate a million integers on the heap, we're going to leverage the Timer class we developed in lesson **8.16 -- Timing your code**. We'll use the Timer class to time how fast our code runs, and show you the performance difference between copying and moving.

```

1  #include <iostream>
2  #include <chrono> // for std::chrono functions
3
4  // Uses the above DynamicArray class
5
6  class Timer
7  {
8  private:
9      // Type aliases to make accessing nested type easier
10     using clock_t = std::chrono::high_resolution_clock;
11     using second_t = std::chrono::duration<double, std::ratio<1> >;
12
13     std::chrono::time_point<clock_t> m_beg;
14
15 public:
16     Timer() : m_beg(clock_t::now())
17     {
18     }
19
20     void reset()
21     {
22         m_beg = clock_t::now();
23     }

```



```

24
25     double elapsed() const
26     {
27         return std::chrono::duration_cast<second_t>(clock_t::now() - m_beg).count();
28     }
29 };
30
31 // Return a copy of arr with all of the values doubled
32 DynamicArray<int> cloneArrayAndDouble(const DynamicArray<int> &arr)
33 {
34     DynamicArray<int> dbl(arr.getLength());
35     for (int i = 0; i < arr.getLength(); ++i)
36         dbl[i] = arr[i] * 2;
37
38     return dbl;
39 }
40
41 int main()
42 {
43     Timer t;
44
45     DynamicArray<int> arr(1000000);
46
47     for (int i = 0; i < arr.getLength(); i++)
48         arr[i] = i;
49
50     arr = cloneArrayAndDouble(arr);
51
52     std::cout << t.elapsed();
53 }

```

On one of the author's machines, in release mode, this program executed in 0.00825559 seconds.

Now let's run the same program again, replacing the copy constructor and copy assignment with a move constructor and move assignment.

```

1  template <class T>
2  class DynamicArray
3  {
4  private:
5      T* m_array;
6      int m_length;
7
8  public:
9      DynamicArray(int length)
10         : m_array(new T[length]), m_length(length)
11     {
12     }
13
14     ~DynamicArray()
15     {
16         delete[] m_array;
17     }
18
19     // Copy constructor
20     DynamicArray(const DynamicArray &arr) = delete;
21
22     // Copy assignment
23     DynamicArray& operator=(const DynamicArray &arr) = delete;
24
25     // Move constructor
26     DynamicArray(DynamicArray &&arr)

```

```

27     : m_length(arr.m_length), m_array(arr.m_array)
28     {
29         arr.m_length = 0;
30         arr.m_array = nullptr;
31     }
32
33     // Move assignment
34     DynamicArray& operator=(DynamicArray &&arr)
35     {
36         if (&arr == this)
37             return *this;
38
39         delete[] m_array;
40
41         m_length = arr.m_length;
42         m_array = arr.m_array;
43         arr.m_length = 0;
44         arr.m_array = nullptr;
45
46         return *this;
47     }
48
49     int getLength() const { return m_length; }
50     T& operator[](int index) { return m_array[index]; }
51     const T& operator[](int index) const { return m_array[index]; }
52
53 };
54
55 #include <iostream>
56 #include <chrono> // for std::chrono functions
57
58 class Timer
59 {
60 private:
61     // Type aliases to make accessing nested type easier
62     using clock_t = std::chrono::high_resolution_clock;
63     using second_t = std::chrono::duration<double, std::ratio<1> >;
64
65     std::chrono::time_point<clock_t> m_beg;
66
67 public:
68     Timer() : m_beg(clock_t::now())
69     {
70     }
71
72     void reset()
73     {
74         m_beg = clock_t::now();
75     }
76
77     double elapsed() const
78     {
79         return std::chrono::duration_cast<second_t>(clock_t::now() - m_beg).count();
80     }
81 };
82
83 // Return a copy of arr with all of the values doubled
84 DynamicArray<int> cloneArrayAndDouble(const DynamicArray<int> &arr)
85 {
86     DynamicArray<int> dbl(arr.getLength());
87     for (int i = 0; i < arr.getLength(); ++i)
88         dbl[i] = arr[i] * 2;

```

```

89
90     return dbl;
91 }
92
93 int main()
94 {
95     Timer t;
96
97     DynamicArray<int> arr(1000000);
98
99     for (int i = 0; i < arr.getLength(); i++)
100         arr[i] = i;
101
102     arr = cloneArrayAndDouble(arr);
103
104     std::cout << t.elapsed();
105 }

```

On the same machine, this program executed in 0.0056 seconds.

Comparing the runtime of the two programs, $0.0056 / 0.00825559 = 67.8\%$. The move version was almost 33% faster!



15.4 -- std::move



Index



15.2 -- R-value references

[C++ TUTORIAL](#) | [PRINT THIS POST](#)

154 comments to 15.3 — Move constructors and move assignment

[« Older Comments](#) [1](#) [2](#)



hellmet

[January 19, 2020 at 4:08 am · Reply](#)

Consider the following code.

```

1  class OBJ {...} // No copies allowed, only moves
2
3  OBJ GetBestObj(/* some params */) {
4      std::map<int, OBJ> obj_map{};
5
6      for (...) {
7          auto obj_inst {OBJ(/* some params */)}
8          obj_map.emplace(obj_inst.GetScore(), obj_inst);
9      }
10

```