

Search:

Not logged in

[Articles](#) [Headers and Includes: Why and How](#)

[register](#)[log in](#)

Published by **Disch**
May 2, 2009 (last update: May 2, 2009)

C++

[Information](#)
[Tutorials](#)
[Reference](#)
[Articles](#)
[Forum](#)

Articles

[Algorithms](#)
[C++ 11](#)
[Graphics and multimedia](#)
[How-To](#)
[Language Features](#)
[Unix/Linux programming](#)
[Source Code](#)
[Standard Library](#)
[Tips and Tricks](#)
[Tools and Libraries](#)
[Visual C++](#)
[Windows API](#)

Headers and Includes: Why and How



Score: 4.6/5 (275 votes)

** 0) Introduction **

This article is meant to address a common newbie problem regarding failure to understand `#include`, headers, and source file interaction. Several good practices are outlined and explained to show how to avoid some ugly snags. Later sections get into more advanced topics (inlining and templates), so even C++ coders with some experience under their belt might benefit from a read!

If you are already familiar with the basics, feel free to skip ahead to section 4. That is where practices and design strategies are discussed.

** 1) Why we need header files. **

If you're just starting out in C++, you might be wondering why you need to `#include` files and why you would want to have multiple `.cpp` files for a program. The reasons for this are simple:

(1) It speeds up compile time. As your program grows, so does your code, and if everything is in a single file, then everything must be fully recompiled every time you make any little change. This might not seem like a big deal for small programs (and it isn't), but when you have a reasonable size project, compile times can take *several minutes* to compile the entire program. Can you imagine having to wait that long between every minor change?

Compile / wait 8 minutes / "oh crap, forgot a semicolon" / compile / wait 8 minutes / debug / compile / wait 8 minutes / etc

(2) It keeps your code more organized. If you separate concepts into specific files, it's easier to find the code you are looking for when you want to make modifications (or just look at it to remember how to use it and/or how it works).

(3) It allows you to separate **interface** from **implementation**. If you don't understand what that means, don't worry, we'll see it in action throughout this article.

Those are the upsides, but the big, obvious downside is that it makes things a little more complicated if you don't understand how it all works (in reality, though, it's simpler than the alternatives as the project gets larger)

C++ programs are built in a two stage process. First, each source file is **compiled** on its own. The compiler generates intermediate files **for each compiled source file**. These intermediate files are often called **object files** -- but they are not to be confused with objects in your code. **Once all the files have been individually compiled, it then links all the object files together, which generates the final binary (the program).**

This means that *each source file is compiled separately* from other source files. As a result of this, in terms of compiling, "a.cpp" is clueless as to what's going on inside of "b.cpp". Here's a quick example to illustrate:

```
1 // in myclass.cpp
2
3 class MyClass
4 {
5 public:
6     void foo();
7     int bar;
8 };
9
10 void MyClass::foo()
11 {
12     // do stuff
13 }
```

```
1 // in main.cpp
2
3 int main()
4 {
5     MyClass a; // Compiler error: 'MyClass' is unidentified
6     return 0;
7 }
```

Even though MyClass is declared *in your program*, it is not declared *in main.cpp*, and therefore when you compile main.cpp you get that error.

This is where header files come in. Header files allow you to make the *interface* (in this case, the class MyClass) visible to other .cpp files, while keeping the *implementation* (in this case, MyClass's member function bodies) in its own .cpp file. That same example again, but tweaked slightly:

```
1 // in myclass.h
2
3 class MyClass
4 {
5 public:
6     void foo();
7     int bar;
8 };
```

```
1 // in myclass.cpp
2 #include "myclass.h"
3
4 void MyClass::foo()
5 {
6 }
```

```

1 //in main.cpp
2 #include "myclass.h" // defines MyClass
3
4 int main()
5 {
6     MyClass a; // no longer produces an error, because MyClass is defined
7     return 0;
8 }

```

The #include statement is basically like a copy/paste operation. The compiler will "replace" the #include line with the actual contents of the file you're including when it **compiles** the file.

```

*****
** 2) What is the difference between .h/.cpp/.hpp/.cc/etc **
*****

```

All files are fundamentally the same in that they're all text files, however different kinds of files should have different extensions:

- *Header files* should use a .h__ extension (.h / .hpp / .hxx). Which of those you use doesn't matter.
- *C++ Source files* should use a .c__ extension (.cpp / .cxx / .cc). Which of those you use doesn't matter.
- *C Source files* should use .c (.c only).

The reason C and C++ source files are separated here is because it makes a difference in some compilers. In all likelihood (since you're reading this on a C++ site), you'll be using C++ code, so refrain from using the .c extension. Also, files with header extensions might be ignored by the compiler if you try to compile them.

So what's **the difference between Header files and Source files**? Basically, header files are #included and not compiled, whereas source files are compiled and not #included. You can try to side-step these conventions and make a file with a source extension behave like a header or vice-versa, but you shouldn't. I won't list the many reasons why you shouldn't (other than the few I already have) -- just don't.

The one exception is that it is sometimes (although *very rarely*) useful to include a source file. This scenario has to do with instantiating templates and is outside the scope of this article. For now... just put it in your brain: "*do not #include source files*" *****

```

** 3) Include guards **
*****

```

C++ compilers do not have brains of their own, and so they will do exactly what you tell them to. If you tell them to include the same file more than once, then that is exactly what they will do. And if you don't handle it properly, you'll get some crazy errors as a result:

```

1 // myclass.h
2
3 class MyClass
4 {
5     void DoSomething() { }
6 };

```

```

1 // main.cpp
2 #include "myclass.h" // define MyClass
3 #include "myclass.h" // Compiler error - MyClass already defined

```

Now you might be saying to yourself "well that's stupid, why would I include the same file twice?" Believe it or not, it will happen a lot. Not quite as illustrated above, though. Usually it happens when you include two files that each include the same file. Example:

```

1 // x.h
2 class X { };

```

```

1 // a.h
2 #include "x.h"
3
4 class A { X x; };

```

```

1 // b.h
2 #include "x.h"
3
4 class B { X x; };

```

```

1 // main.cpp
2
3 #include "a.h" // also includes "x.h"
4 #include "b.h" // includes x.h again! ERROR

```

Because of this scenario, many people are told not to put `#include` in header files. However *this is bad advice and you should not listen to it*. Sadly, some people are actually *taught* this in C++ courses that they are *paying money for*. If your C++ instructor tells you not to `#include` in header files, then [grudgingly] follow his instructions in order to pass the course, but once you're out of his/her class, shake that habit.

The truth is there is nothing wrong with putting `#include` in header files -- and in fact it is very beneficial. *Provided you take two precautions:*

- 1) Only `#include` things you *need to include* (covered next section)
- 2) `Guard` against incidental multiple includes with include guards.

An Include Guard is a technique which uses a unique identifier that you `#define` at the top of the file. Here's an example:

```

1 //x.h
2
3 #ifndef __X_H_INCLUDED__ // if x.h hasn't been included yet...
4 #define __X_H_INCLUDED__ // #define this so the compiler knows it has been included
5
6 class X { };

```

```
7 |
8 | #endif
```

This works by skipping over the entire header if it was already included. `__X_H_INCLUDED__` is #defined the first time `x.h` is included -- and if `x.h` is included a second time, the compiler will skip over the header because the `#ifndef` check will fail.

Always guard your headers. Always always always. It doesn't hurt anything to do it, and it will save you some headaches. For the rest of this article, it is assumed all header files are include guarded (even if I don't explicitly put it in the example).

You do not need to guard your .cpp files, because they are not #included (Or at least they shouldn't be... right? *RIGHT?*)

```
*****
** 4) The "right way" to include **
*****
```

Classes you create will often have dependencies on other classes. A derived class, for example, will always be dependent on its parent, because in order to be derived from the parent, it must be aware of its parent at compile time.

There are **two basic kinds of dependencies** you need to be aware of:

- 1) stuff that can be **forward declared**
- 2) stuff that needs to be **#included**

If, for example, class A uses class B, then class B is one of class A's dependencies. Whether it can be **forward declared** or needs to be **included** depends on how B is used within A:

- **do nothing** if: A makes no references at all to B
- **do nothing** if: The only reference to B is in a **friend declaration**
- **forward declare B** if: A contains a **B pointer or reference**: `B* myb;`
- **forward declare B** if: **one or more functions has a B object/pointer/reference as a parameter, or as a return type**: `B MyFunction(B myb);`
- **#include "b.h"** if: B is a parent class of A
- **#include "b.h"** if: A contains a B object: `B myb;`

You want to do the least drastic option possible. Do nothing if you can, but if you can't, forward declare if you can. But if it's necessary, then `#include` the other header.

Ideally, **the dependencies for the class should be layed out in the header**. Here is a typical outline of how a "right way" header might look:

`myclass.h`

```
1 //=====
2 // include guard
3 #ifndef __MYCLASS_H_INCLUDED__
4 #define __MYCLASS_H_INCLUDED__
5
```

```

6 //=====
7 // forward declared dependencies
8 class Foo;
9 class Bar;
10
11 //=====
12 // included dependencies
13 #include <vector>
14 #include "parent.h"
15
16 //=====
17 // the actual class
18 class MyClass : public Parent // Parent object, so #include "parent.h"
19 {
20 public:
21     std::vector<int> avector; // vector object, so #include <vector>
22     Foo* foo; // Foo pointer, so forward declare Foo
23     void Func(Bar& bar); // Bar reference, so forward declare Bar
24
25     friend class MyFriend; // friend declaration is not a dependency
26                          // don't do anything about MyFriend
27 };
28
29 #endif // __MYCLASS_H_INCLUDED__

```

This shows the two different kinds of dependencies and how to handle them. Because MyClass only uses a pointer to Foo and not a full Foo object, we can forward declare Foo, and don't need to #include "foo.h". *You should always forward declare what you can -- don't #include unless it's necessary. Unnecessary #includes can lead to trouble.*

If you stick to this system, you will bulletproof yourself, and will minimize #include related hazards.

```

*****
** 5) Why that is the "right way" to include **
*****

```

Note: in this section I refer to the "right way" method outlined above as "mine". While I did come up with it on my own (after struggling through the mucky muck for a while) -- I can't say I was the first person who ever thought of it, so it isn't really "mine". But for purposes of this article, *I call it "mine" for simplicity.*

You: "So-and-so" says that #including in a header is dangerous, but you say it's not! Why is your way so much better than what so-and-so says?"

So-and-so is partially right, but is explaining it wrong. Frivolous and careless #includes *can* lead to trouble. And one way to avoid those troubles *is* to never #include inside a header file. So yeah, so-and-so's heart is in the right place. But ultimately, using so-and-so's approach is going to give yourself TONS of additional work and headaches.

The concept I'm illustrating is very OO, and enhances encapsulation. *The general idea is that it makes "myclass.h" fully self-contained and doesn't require any other area of the program* (other than MyClass's implementation/source file) to know how MyClass works internally. If some other class needs to use MyClass, it can just #include "myclass.h" and be done with it!

The alternative (so-and-so's method), would require you to `#include` all of `MyClass`'s dependencies *before* `#including` `"myclass.h"`, since `myclass.h` can't include its dependencies itself. This is headache city, because now using a class isn't so straightforward.

Here is an example of why my method is good:

```
1 //example.cpp
2
3 // I want to use MyClass
4 #include "myclass.h" // will always work, no matter what MyClass looks like.
5                     // You're done
6                     // (provided myclass.h follows my outline above and does
7                     // not make unnecessary #includes)
```

Here is an example of why so-and-so's method is bad:

```
1 //example.cpp
2
3 // I want to use MyClass
4 #include "myclass.h"
5 // ERROR 'Parent' undefined
```

so-and-so: "Hrm... okay...."

```
1 #include "parent.h"
2 #include "myclass.h"
3 // ERROR 'std::vector' undefined
```

```
1 #include "parent.h"
2 #include <vector>
3 #include "myclass.h"
4 // ERROR 'Support' undefined
```

so-and-so: "WTF? MyClass doesn't even use Support! But alright..."

```
1 #include "parent.h"
2 #include <vector>
3 #include "support.h"
4 #include "myclass.h"
5 // ERROR 'Support' undefined
```

so-and-so: "Give me a break! I'm including it! What else do you want!"

Believe it or not, the above *does* happen. Little did poor so-and-so know, but `"parent.h"` uses `Support`, and therefore you must `#include "support.h"` *before* `"parent.h"`.

And what happens if support.h needs something else? What if *that something else* needs something else? We're already up to 4 #includes just to use a single class! With so-and-so's method, not only do you have to remember which includes are needed for each class, but also *the order in which you need to #include them*. This becomes a *huge* nightmare *very* quickly.


And what happens if you want to tweak MyClass? Let's say you decide that it would be better to use std::list instead of std::vector. With so-and-so's method, you now have to go back and change *every single file* that #includes "myclass.h" and change it to include <list> instead of <vector> (which might be dozens of files depending on the size of the project and how often MyClass is used), whereas with my method you only have to change "myclass.h", and maybe "myclass.cpp".

The "right way" I illustrated above is all about **encapsulation**. Files that want to use MyClass don't need to be aware of what MyClass uses in order for it to work, and don't need to #include any MyClass dependencies. All you need to do to get MyClass to work is #include "myclass.h". Period. **The header file is set up to be completely self contained**. It's all very OO friendly, very easy to use, and very easy to maintain.

** 6) Circular Dependencies **

A circular dependency is when two (or more) classes depend on each other. For example, class A depends on class B, and class B depends on class A. *If you stick to "the right way" and forward declare when you can instead of #including* needlessly, this usually isn't a problem. As long as **the circle is broken** with a **forward declaration** at some point, you're fine.

Here's the perfect example of why you should only #include what is necessary:

```
1 // a.h -- assume it's guarded
2 #include "b.h"
3
4 class A { B* b; }; 
```

```
1 // b.h -- assume it's guarded
2 #include "a.h"
3
4 class B { A* a };
```

An initial glance might see nothing wrong with this. B is a dependency of A, so you include it, and A is a dependency of B, so you include it. **So what's wrong with this?**

This is a circular inclusion (also called an infinite inclusion) and is the result of one or more includes that shouldn't be there. Say for example you compile "a.cpp":

```
1 // a.cpp
2 #include "a.h"
```


The compiler will do the following:

```

1 #include "a.h"
2
3 // start compiling a.h
4 #include "b.h"
5
6 // start compiling b.h
7 #include "a.h"
8
9 // compilation of a.h skipped because it's guarded
10
11 // resume compiling b.h
12 class B { A* a }; // <--- ERROR, A is undeclared

```

Even though you're #including "a.h", the compiler is not seeing the A class until after the B class gets compiled. This is because of the circular inclusion problem. This is why you should **always forward declare** when you're only using a pointer or reference. Here, "a.h" should not be #including b.h, but instead should just be forward declaring B. Likewise, b.h should be forward declaring A. **If you make those changes, the problem is solved.**

The circular inclusion problem may persist if both dependencies are #include dependencies (ie: they can't be forward declared). Here's an example:

```

1 // a.h (guarded)
2
3 #include "b.h"
4
5 class A
6 {
7     B b; // B is an object, can't be forward declared
8 };

```

```

1 // b.h (guarded)
2
3 #include "a.h"
4
5 class B
6 {
7     A a; // A is an object, can't be forward declared
8 };

```

You may note, however, that this situation is *conceptually impossible*. There is a fundamental design flaw. If A has a B object, and B has an A object, then A contains a B, which contains another A, which contains another B, which contains another A, which contains another B, etc, etc. You have an infinite recursion problem, and either class is simply impossible to instantiate. **The solution is to have one or both classes contain a *pointer or reference* to the other, rather than a full object. Then you can forward declare, and then you can get around the circular inclusion problem.**

** 7) Function inlining **

The catch-22 with inline functions is that their **function body** needs to exist in every cpp file which calls them, otherwise you get linker errors (since they can't be linked during the linker process -- they need to be compiled into the code during the compiler process).

This might open circular reference holes or other scenarios that might complicate the "right way" outline.

```

1 class B
2 {
3 public:
4     void Func(const A& a)    // parameter, so forward declare is okay
5     {
6         a.DoSomething();    // but now that we've dereferenced it, it
7                             // becomes an #include dependency
8                             // = we now have a potential circular inclusion
9     }
10 };

```

The key is that while **inline function need to exist in the header**, they *do not* need to exist in **the class definition** itself. This lets us exploit a loophole:

```

1 // b.h (assume its guarded)
2
3 //-----
4 class A; // forward declared dependency
5
6 //-----
7 class B
8 {
9 public:
10     void Func(const A& a); // okay, A is forward declared
11 };
12
13 //-----
14 #include "a.h" // A is now an include dependency
15
16 inline void B::Func(const A& a)
17 {
18     a.DoSomething(); // okay! a.h has been included
19 }

```

While you might not think so at first glance... this is *perfectly safe*. The circular inclusion problem is avoided completely, even if a.h includes b.h. This is because the additional #includes don't come up until AFTER class B is fully defined, and they are therefore harmless.

You: "But putting function bodies at the end of my header is ugly. Isn't there a way to avoid that?"

Me: Yep! Just move the function bodies to another header:

```

1 // b.h
2
3     // blah blah
4
5 class B { /* blah blah */ };
6
7 #include "b_inline.h" // or I sometimes use "b.hpp"

```

```

1 // b_inline.h (or b.hpp -- whatever)
2
3 #include "a.h"
4 #include "b.h" // not necessary, but harmless
5                // you can do this to make this "feel" like a source
6                // file, even though it isn't
7
8 inline void B::Func(const A& a)
9 {
10     a.DoSomething();
11 }

```

This separates the interface from the implementation, while still allowing the implementation to be inlined. You can also have a normal "b.cpp" file for the implementation that isn't inlined.

** 8) Forward declaring templates **

Forward declaring is pretty straight-forward when it comes to simple classes, but when dealing with template classes, things aren't so simple. Consider the following scenario:

```

1 // a.h
2
3 // included dependencies
4 #include "b.h"
5
6 // the class template
7 template <typename T>
8 class Tem
9 {
10     /*...*/
11     B b;
12 };
13
14 // class most commonly used with 'int'
15 typedef Tem<int> A; // typedef'd as 'A'

```

```

1 // b.h
2
3 // forward declared dependencies
4 class A; // error!
5
6 // the class
7 class B
8 {
9     /* ... */
10    A* ptr;
11 };

```

While this seems perfectly logical, it doesn't work! (Although, logically you really think it should. This is an irritation of the language). Because 'A' isn't really a class, but rather a typedef, the compiler will bark at you. Also notice that we can't just #include "a.h" here because of a circular dependency problem.

In order to forward declare 'A', we need to typedef it. Which means we need to forward declare what it's typedef'd as. The way to forward declare it is like so:

```

1 template <typename T> class Tem; // forward declare our template
2 typedef Tem<int> A; // then typedef 'A'

```

This is quite a bit uglier than `class A;`, but nonetheless is a necessary evil. This, however, makes templated classes hard to encapsulate, because it requires every class which forward declares them to know exactly how the template is layed out. If that ever changes you have a big mess to clean up.

A practical solution to this problem is to create an alternative header which has the forward declarations of your templated classes and their typedefs. Here's a more elegant way to approach the above example:

```

1 //a.h
2
3 #include "b.h"
4
5 template <typename T>
6 class Tem
7 {
8     /*...*/
9     B b;
10 };

```

```

1 //a_fwd.h
2
3 template <typename T> class Tem;
4 typedef Tem<int> A;

```

```

1 //b.h
2

```

```
3 #include "a_fwd.h"
4
5 class B
6 {
7     /*...*/
8     A* ptr;
9 };
```

This allows B to include a header which **forward declares A** without including **the entire class definition**.

Rate this

Please, choose stars to award:



(hover mouse and click)

[Home page](#) | [Privacy policy](#)

© cplusplus.com, 2000-2017 - All rights reserved - v3.1

[Spotted an error? contact us](#)