[C++项目中的extern "C" {}](http://www.cnblogs.com/skynet/archive/2010/07/10/1774964.html)

2010-07-10 19:45 by 吴秦, 49293 阅读, 18 评论, [收藏](http://www.cnblogs.com/skynet/archive/2010/07/10/1774964.html), [编辑](https://i.cnblogs.com/EditPosts.aspx?postid=1774964)

**引言**

在用C++的项目源码中，经常会不可避免的会看到下面的代码：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9 | #ifdef \_\_cplusplus  extern "C" {  #endif    /\*...\*/    #ifdef \_\_cplusplus  }  #endif |

它到底有什么用呢，你知道吗？而且这样的问题经常会出现在面试or笔试中。下面我就从以下几个方面来介绍它：

* 1、#ifdef \_cplusplus/#endif \_cplusplus及发散
* 2、extern "C"
  + 2.1、extern关键字
  + 2.2、"C"
  + 2.3、小结extern "C"
* 3、C和C++互相调用
  + 3.1、C++的编译和连接
  + 3.2、C的编译和连接
  + 3.3、C++中调用C的代码
  + 3.4、C中调用C++的代码
* 4、C和C++混合调用特别之处函数指针

**1、#ifdef \_cplusplus/#endif \_cplusplus及发散**

在介绍extern "C"之前，我们来看下#ifdef \_cplusplus/#endif \_cplusplus的作用。很明显#ifdef/#endif、#ifndef/#endif用于条件编译，#ifdef \_cplusplus/#endif \_cplusplus——表示如果定义了宏\_cplusplus，就执行#ifdef/#endif之间的语句，否则就不执行。

在这里为什么需要#ifdef \_cplusplus/#endif \_cplusplus呢？因为C语言中不支持extern "C"声明，如果你明白extern "C"的作用就知道在C中也没有必要这样做，这就是条件编译的作用！在.c文件中包含了extern "C"时会出现编译时错误。

既然说到了条件编译，我就介绍它的一个重要应用——**避免重复包含头文件**。还记得腾讯笔试就考过这个题目，给出类似下面的代码（下面是我最近在研究的一个开源web服务器——Mongoose的头文件mongoose.h中的一段代码）：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17 | #ifndef MONGOOSE\_HEADER\_INCLUDED  #define    MONGOOSE\_HEADER\_INCLUDED    #ifdef \_\_cplusplus  extern "C" {  #endif /\* \_\_cplusplus \*/    /\*.................................   \* do something here   \*.................................   \*/    #ifdef \_\_cplusplus  }  #endif /\* \_\_cplusplus \*/    #endif /\* MONGOOSE\_HEADER\_INCLUDED \*/ |

然后叫你说明上面宏#ifndef/#endif的作用？为了解释一个问题，我们先来看两个事实：

* 这个头文件mongoose.h可能在项目中被多个源文件包含（#include "mongoose.h"），而对于一个大型项目来说，这些冗余可能导致错误，因为一个头文件包含类定义或inline函数，在一个源文件中mongoose.h可能会被#include两次（如，a.h头文件包含了mongoose.h，而在b.c文件中#include a.h和mongoose.h）——这就会出错（在同一个源文件中一个结构体、类等被定义了两次）。
* 从逻辑观点和减少编译时间上，都要求去除这些冗余。然而让程序员去分析和去掉这些冗余，不仅枯燥且不太实际，**最重要的是有时候又需要这种冗余来保证各个模块的独立**。

为了解决这个问题，上面代码中的

#ifndef MONGOOSE\_HEADER\_INCLUDED   
#define    MONGOOSE\_HEADER\_INCLUDED   
/\*……………………………\*/   
#endif /\* MONGOOSE\_HEADER\_INCLUDED \*/

就起作用了。如果定义了MONGOOSE\_HEADER\_INCLUDED，#ifndef/#endif之间的内容就被忽略掉。因此，编译时第一次看到mongoose.h头文件，它的内容会被读取且给定MONGOOSE\_HEADER\_INCLUDED一个值。之后再次看到mongoose.h头文件时，MONGOOSE\_HEADER\_INCLUDED就已经定义了，mongoose.h的内容就不会再次被读取了。

**2、extern "C"**

首先从字面上分析extern "C"，它由两部分组成——extern关键字、"C"。下面我就从这两个方面来解读extern "C"的含义。

**2.1、extern关键字**

在一个项目中必须保证函数、变量、枚举等在所有的源文件中保持一致，除非你指定定义为局部的。首先来一个例子：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7 | //file1.c:      int x=1;      int f(){do something here}  //file2.c:      extern int x;      int f();      void g(){x=f();} |

在file2.c中g()使用的x和f()是定义在file1.c中的。extern关键字表明file2.c中x，仅仅是一个变量的声明，其并不是在定义变量x，并未为x分配内存空间。变量x在所有模块中作为一种全局变量只能被定义一次，否则会出现连接错误。但是可以声明多次，且声明必须保证类型一致，如：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9 | //file1.c:      int x=1;      int b=1;      extern c;  //file2.c:      int x;// x equals to default of int type 0      int f();      extern double b;      extern int c; |

在这段代码中存在着这样的三个错误：

1. x被定义了两次
2. b两次被声明为不同的类型
3. c被声明了两次，但却没有定义

回到extern关键字，extern是C/C++语言中表明**函数**和**全局变量**作用范围（可见性）的关键字，该关键字告诉编译器，其声明的函数和变量可以在本模块或其它模块中使用。通常，在模块的头文件中对本模块提供给其它模块引用的函数和全局变量以关键字extern声明。例如，如果模块B欲引用该模块A中定义的全局变量和函数时只需包含模块A的头文件即可。这样，模块B中调用模块A中的函数时，在编译阶段，模块B虽然找不到该函数，但是并不会报错；它会在连接阶段中从模块A编译生成的目标代码中找到此函数。

与extern对应的关键字是 static，被它修饰的全局变量和函数只能在本模块中使用。因此，一个函数或变量只可能被本模块使用时，其不可能被extern “C”修饰。

**2.2、"C"**

典型的，一个C++程序包含其它语言编写的部分代码。类似的，C++编写的代码片段可能被使用在其它语言编写的代码中。不同语言编写的代码互相调用是困难的，甚至是同一种编写的代码但不同的编译器编译的代码。例如，不同语言和同种语言的不同实现可能会在注册变量保持参数和参数在栈上的布局，这个方面不一样。

为了使它们遵守统一规则，可以使用extern指定一个编译和连接规约。例如，声明C和C++标准库函数strcyp()，并指定它应该根据C的编译和连接规约来链接：

|  |  |
| --- | --- |
| 1 | extern "C" char\* strcpy(char\*,const char\*); |

注意它与下面的声明的不同之处：

|  |  |
| --- | --- |
| 1 | extern char\* strcpy(char\*,const char\*); |

下面的这个声明仅表示在连接的时候调用strcpy()。

extern "C"指令非常有用，因为C和C++的近亲关系。**注意：extern "C"指令中的C，表示的一种编译和连接规约，而不是一种语言。C表示符合C语言的编译和连接规约的任何语言，如Fortran、assembler等。**

还有要说明的是，extern "C"指令仅指定编译和连接规约，但不影响语义。例如在函数声明中，指定了extern "C"，仍然要遵守C++的类型检测、参数转换规则。

再看下面的一个例子，为了声明一个变量而不是定义一个变量，你必须在声明时指定extern关键字，但是当你又加上了"C"，它不会改变语义，但是会改变它的编译和连接方式。

如果你有很多语言要加上extern "C"，你可以将它们放到extern "C"{ }中。

**2.3、小结extern "C"**

通过上面两节的分析，我们知道extern "C"的真实目的是实现**类C和C++的混合编程**。在C++源文件中的语句前面加上extern "C"，表明它按照类C的编译和连接规约来编译和连接，而不是C++的编译的连接规约。这样在类C的代码中就可以调用C++的函数or变量等。（注：我在这里所说的类C，代表的是跟C语言的编译和连接方式一致的所有语言）

**3、C和C++互相调用**

我们既然知道extern "C"是实现的类C和C++的混合编程。下面我们就分别介绍如何在C++中调用C的代码、C中调用C++的代码。首先要明白C和C++互相调用，你得知道它们之间的编译和连接差异，及如何利用extern "C"来实现相互调用。

**3.1、C++的编译和连接**

C++是一个面向对象语言（虽不是纯粹的面向对象语言），它支持函数的重载，重载这个特性给我们带来了很大的便利。为了支持函数重载的这个特性，C++编译器实际上将下面这些重载函数：

|  |  |
| --- | --- |
| 1  2  3  4 | void print(int i);  void print(char c);  void print(float f);  void print(char\* s); |

编译为：

|  |  |
| --- | --- |
| 1  2  3  4 | \_print\_int  \_print\_char  \_print\_float  \_pirnt\_string |

这样的函数名，来唯一标识每个函数。注：不同的编译器实现可能不一样，但是都是利用这种机制。所以当连接是调用print(3)时，它会去查找\_print\_int(3)这样的函数。下面说个题外话，正是因为这点，重载被认为不是多态，多态是运行时动态绑定（“一种接口多种实现”），如果硬要认为重载是多态，它顶多是编译时“多态”。

C++中的变量，编译也类似，如全局变量可能编译g\_xx，类变量编译为c\_xx等。连接是也是按照这种机制去查找相应的变量。

**3.2、C的编译和连接**

C语言中并没有重载和类这些特性，故并不像C++那样print(int i)，会被编译为\_print\_int，而是直接编译为\_print等。因此如果直接在C++中调用C的函数会失败，因为连接是调用C中的print(3)时，它会去找\_print\_int(3)。因此extern "C"的作用就体现出来了。

**3.3、C++中调用C的代码**

假设一个C的头文件cHeader.h中包含一个函数print(int i)，为了在C++中能够调用它，必须要加上extern关键字（原因在extern关键字那节已经介绍）。它的代码如下：

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | #ifndef C\_HEADER  #define C\_HEADER    extern void print(int i);    #endif C\_HEADER |

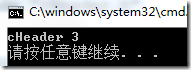
相对应的实现文件为cHeader.c的代码为：

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | #include <stdio.h>  #include "cHeader.h"  void print(int i)  {      printf("cHeader %d\n",i);  } |

现在C++的代码文件C++.cpp中引用C中的print(int i)函数：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9 | extern "C"{  #include "cHeader.h"  }    int main(int argc,char\*\* argv)  {      print(3);      return 0;  } |

执行程序输出：

[](http://images.cnblogs.com/cnblogs_com/skynet/WindowsLiveWriter/externCinternals_16B1/image_2.png)

**3.4、C中调用C++的代码**

现在换成在C中调用C++的代码，这与在C++中调用C的代码有所不同。如下在cppHeader.h头文件中定义了下面的代码：

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | #ifndef CPP\_HEADER  #define CPP\_HEADER    extern "C" void print(int i);    #endif CPP\_HEADER |

相应的实现文件cppHeader.cpp文件中代码如下：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8 | #include "cppHeader.h"    #include <iostream>  using namespace std;  void print(int i)  {      cout<<"cppHeader "<<i<<endl;  } |

在C的代码文件c.c中调用print函数：

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | extern void print(int i);  int main(int argc,char\*\* argv)  {      print(3);      return 0;  } |

注意在C的代码文件中直接#include "cppHeader.h"头文件，编译出错。而且如果不加extern int print(int i)编译也会出错。

**4、C和C++混合调用特别之处函数指针**

当我们C和C++混合编程时，有时候会用一种语言定义函数指针，而在应用中将函数指针指向另一中语言定义的函数。如果C和C++共享同一中编译和连接、函数调用机制，这样做是可以的。然而，这样的通用机制，通常不然假定它存在，因此我们必须小心地确保函数以期望的方式调用。

而且当指定一个函数指针的编译和连接方式时，函数的所有类型，包括函数名、函数引入的变量也按照指定的方式编译和连接。如下例：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28 | typedef int (\*FT) (const void\* ,const void\*);//style of C++    extern "C"{      typedef int (\*CFT) (const void\*,const void\*);//style of C      void qsort(void\* p,size\_t n,size\_t sz,CFT cmp);//style of C  }    void isort(void\* p,size\_t n,size\_t sz,FT cmp);//style of C++  void xsort(void\* p,size\_t n,size\_t sz,CFT cmp);//style of C    //style of C  extern "C" void ysort(void\* p,size\_t n,size\_t sz,FT cmp);    int compare(const void\*,const void\*);//style of C++  extern "C" ccomp(const void\*,const void\*);//style of C    void f(char\* v,int sz)  {      //error,as qsort is style of C      //but compare is style of C++      qsort(v,sz,1,&compare);      qsort(v,sz,1,&ccomp);//ok        isort(v,sz,1,&compare);//ok      //error,as isort is style of C++      //but ccomp is style of C      isort(v,sz,1,&ccopm);  } |

**注意：typedef int (\*FT) (const void\* ,const void\*)，表示定义了一个函数指针的别名FT，这种函数指针指向的函数有这样的特征：返回值为int型、有两个参数，参数类型可以为任意类型的指针（因为为void\*）。**

最典型的函数指针的别名的例子是，信号处理函数signal，它的定义如下：

|  |  |
| --- | --- |
| 1  2 | typedef void (\*HANDLER)(int);  HANDLER signal(int ,HANDLER); |

上面的代码定义了信函处理函数signal，它的返回值类型为HANDLER，有两个参数分别为int、HANDLER。 这样避免了要这样定义signal函数：

|  |  |
| --- | --- |
| 1 | void (\*signal (int ,void(\*)(int) ))(int) |

比较之后可以明显的体会到typedef的好处。