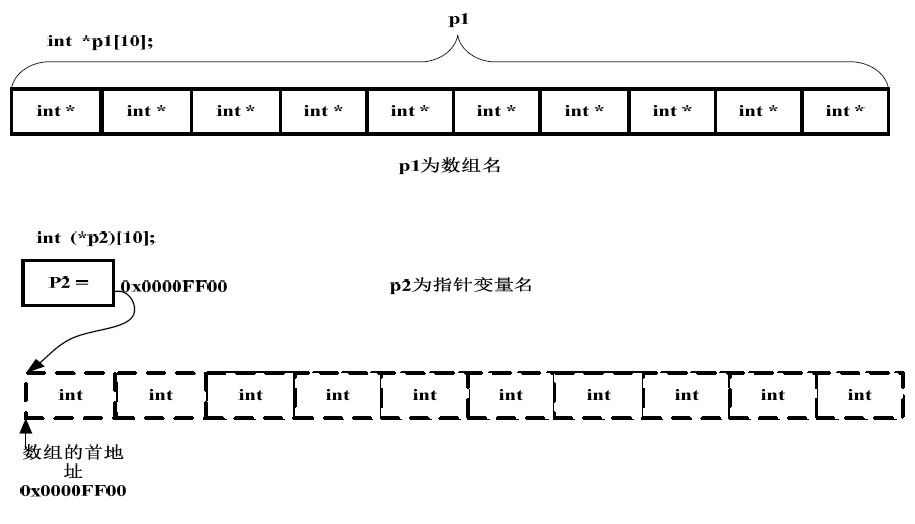
<http://c.biancheng.net/cpp/html/476.html>

**一、指针数组和数组指针的内存布局**

初学者总是分不出指针数组与数组指针的区别。其实很好理解：  
指针数组：首先它是一个数组，数组的元素都是指针，数组占多少个字节由数组本身决定。它是“储存指针的数组”的简称。  
数组指针：首先它是一个指针，它指向一个数组。在32 位系统下永远是占4 个字节，至于它指向的数组占多少字节，不知道。它是“指向数组的指针”的简称。  
  
下面到底哪个是数组指针，哪个是指针数组呢：  
A)  
int \*p1[10];  
B)  
int (\*p2)[10];  
每次上课问这个问题，总有弄不清楚的。这里需要明白一个符号之间的优先级问题。  
  
“[]”的优先级比“\*”要高。p1 先与“[]”结合，构成一个数组的定义，数组名为p1，int \*修饰的是数组的内容，即数组的每个元素。那现在我们清楚，这是一个数组，其包含10 个指向int 类型数据的指针，即指针数组。至于p2 就更好理解了，在这里“（）”的优先级比“[]”高，“\*”号和p2 构成一个指针的定义，指针变量名为p2，int 修饰的是数组的内容，即数组的每个元素。数组在这里并没有名字，是个匿名数组。那现在我们清楚p2 是一个指针，它指向一个包含10 个int 类型数据的数组，即数组指针。我们可以借助下面的图加深理解：



**二、int (\*)[10] p2-----也许应该这么定义数组指针**

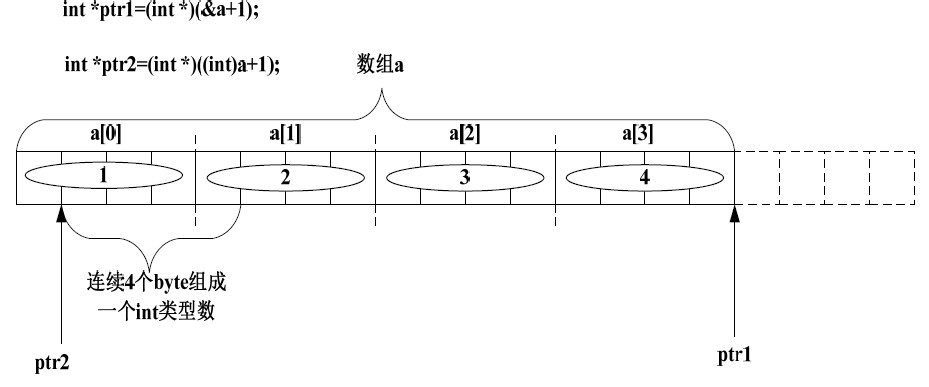
这里有个有意思的话题值得探讨一下：平时我们定义指针不都是在数据类型后面加上指针变量名么？这个指针p2 的定义怎么不是按照这个语法来定义的呢？也许我们应该这样来定义p2：  
   int (\*)[10] p2;  
int (\*)[10]是指针类型，p2 是指针变量。这样看起来的确不错，不过就是样子有些别扭。其实数组指针的原型确实就是这样子的，只不过为了方便与好看把指针变量p2 前移了而已。你私下完全可以这么理解这点。虽然编译器不这么想。^\_^

**三、再论a 和&a 之间的区别**

既然这样，那问题就来了。前面我们讲过a 和&a 之间的区别，现在再来看看下面的代码：  
int main()  
{  
   char a[5]={'A','B','C','D'};  
   char (\*p3)[5] = &a;  
   char (\*p4)[5] = a;  
   return 0;  
}  
上面对p3 和p4 的使用，哪个正确呢？p3+1 的值会是什么？p4+1 的值又会是什么？毫无疑问，p3 和p4 都是数组指针，指向的是整个数组。&a 是整个数组的首地址，a是数组首元素的首地址，其值相同但意义不同。在C 语言里，赋值符号“=”号两边的数据类型必须是相同的，如果不同需要显示或隐式的类型转换。p3 这个定义的“=”号两边的数据类型完全一致，而p4 这个定义的“=”号两边的数据类型就不一致了。左边的类型是指向整个数组的指针，右边的数据类型是指向单个字符的指针。在Visual C++6.0 上给出如下警告：  
   warning C4047: 'initializing' : 'char (\*)[5]' differs in levels of indirection from 'char \*'。  
还好，这里虽然给出了警告，但由于&a 和a 的值一样，而变量作为右值时编译器只是取变量的值，所以运行并没有什么问题。不过我仍然警告你别这么用。  
  
既然现在清楚了p3 和p4 都是指向整个数组的，那p3+1 和p4+1 的值就很好理解了。  
  
但是如果修改一下代码，会有什么问题？p3+1 和p4+1 的值又是多少呢？  
int main()  
{  
   char a[5]={'A','B','C','D'};  
   char (\*p3)[3] = &a;  
   char (\*p4)[3] = a;  
   return 0;  
}  
  
甚至还可以把代码再修改：  
int main()  
{  
   char a[5]={'A','B','C','D'};  
   char (\*p3)[10] = &a;  
   char (\*p4)[10] = a;  
   return 0;  
}  
这个时候又会有什么样的问题？p3+1 和p4+1 的值又是多少？  
  
上述几个问题，希望读者能仔细考虑考虑。

**四、地址的强制转换**

先看下面这个例子：  
struct Test  
{  
   int Num;  
   char \*pcName;  
   short sDate;  
   char cha[2];  
   short sBa[4];  
}\*p;  
  
假设p 的值为0x100000。如下表表达式的值分别为多少？  
   p + 0x1 = 0x\_\_\_ ?  
   (unsigned long)p + 0x1 = 0x\_\_\_?  
   (unsigned int\*)p + 0x1 = 0x\_\_\_?  
我相信会有很多人一开始没看明白这个问题是什么意思。其实我们再仔细看看，这个知识点似曾相识。一个指针变量与一个整数相加减，到底该怎么解析呢？  
  
还记得前面我们的表达式“a+1”与“&a+1”之间的区别吗？其实这里也一样。指针变量与一个整数相加减并不是用指针变量里的地址直接加减这个整数。这个整数的单位不是byte 而是元素的个数。所以：p + 0x1 的值为0x100000+sizof（Test）\*0x1。至于此结构体的大小为20byte，前面的章节已经详细讲解过。所以p +0x1 的值为：0x100014。  
  
(unsigned long)p + 0x1 的值呢？这里涉及到强制转换，将指针变量p 保存的值强制转换成无符号的长整型数。任何数值一旦被强制转换，其类型就改变了。所以这个表达式其实就是一个无符号的长整型数加上另一个整数。所以其值为：0x100001。  
  
(unsigned int\*)p + 0x1 的值呢？这里的p 被强制转换成一个指向无符号整型的指针。所以其值为：0x100000+sizof（unsigned int）\*0x1，等于0x100004。  
  
上面这个问题似乎还没啥技术含量，下面就来个有技术含量的：在x86 系统下，其值为多少？  
intmain()  
{  
   int a[4]={1,2,3,4};  
   int \*ptr1=(int \*)(&a+1);  
   int \*ptr2=(int \*)((int)a+1);  
   printf("%x,%x",ptr1[-1],\*ptr2);  
   return 0;  
}  
这是我讲课时一个学生问我的题，他在网上看到的，据说难倒了n 个人。我看题之后告诉他，这些人肯定不懂汇编，一个懂汇编的人，这种题实在是小case。下面就来分析分析这个问题：  
  
根据上面的讲解，&a+1 与a+1 的区别已经清楚。  
  
ptr1：将&a+1 的值强制转换成int\*类型，赋值给int\* 类型的变量ptr，ptr1 肯定指到数组a 的下一个int 类型数据了。ptr1[-1]被解析成\*(ptr1-1)，即ptr1 往后退4 个byte。所以其值为0x4。  
ptr2：按照上面的讲解，(int)a+1 的值是元素a[0]的第二个字节的地址。然后把这个地址强制转换成int\*类型的值赋给ptr2，也就是说\*ptr2 的值应该为元素a[0]的第二个字节开始的连续4 个byte 的内容。  
  
其内存布局如下图：



好，问题就来了，这连续4 个byte 里到底存了什么东西呢？也就是说元素a[0],a[1]里面的值到底怎么存储的。这就涉及到系统的大小端模式了，如果懂汇编的话，这根本就不是问题。既然不知道当前系统是什么模式，那就得想办法测试。大小端模式与测试的方法在第一章讲解union 关键字时已经详细讨论过了，请翻到彼处参看，这里就不再详述。我们可以用下面这个函数来测试当前系统的模式。  
int checkSystem( )  
{  
   union check  
   {  
      int i;  
     char ch;  
   } c;  
   c.i = 1;  
   return (c.ch ==1);  
}  
如果当前系统为大端模式这个函数返回0；如果为小端模式，函数返回1。也就是说如果此函数的返回值为1 的话，\*ptr2 的值为0x2000000。如果此函数的返回值为0 的话，\*ptr2 的值为0x100。