**字符、标识符及关键字**

**1. 字符集**

C语言源程序由字符序列构成，C语言字符集是7位ASCII码的子集，组成如下：

英文字母：a~z和A~Z

数字字符：0~9

特殊字符：! “ # % & ‘ ( ) \* + , - . / : < > = ? [ ] \ ^ \_ { } | ~

空白字符：空格，换行符，水平制表符，垂直制表符，换页符

**2. 词法元素**

编译器可对C程序中的字符序列进行词法分析，即按照特定规则分解成称之为记号的词法元素。记号是程序中具有语义的最基本组成单位，共分为5类：标识符、关键字、常量、运算符和标点符号。编译器从左至右收集字符，并总是尽量建立最长的记号，即使结果并不构成有效的C语言程序。相邻记号可以用空白符或注释语句分开。

【示例】 x+++++y（非法），x++ + ++y（合法）

对于x+++++y，编译器会把这些字符分解成x、++、++、+、y共5个记号，这是无效的C语法。如组合成记号后的x++、+、++y虽是有效的语法，但编译器不会分解为这种记号组合。

**3. 标识符**

标识符是用来标识用户定义的常量、变量、数据类型和函数等名字的符号。C语言规定，标识符由字母、数字和下划线组成，但首字符必须是字母或下划线，字母要区分大小写。

下面是C语言中一些合法的标识符：

K \_id x\_coord time1

但以下的表示均不是C语言中的合法标识符：

20\_sum （不能以数字开头）

not#me （特殊符号#不是组成字符）

创建标识符是为了对C程序中的对象进行唯一地命名。不能使用类似double和if这样的C语言关键字为自己程序的对象命名，同时也要避免使用C语言程序库中的函数名和常量名（如scanf和sin）为自己程序的对象命名。

**4. 关键字**

关键字也称作保留字，是被系统赋予特定含义并有专门用途的标识符。关键字不能作为普通标识符，但可以作为宏名，因为所有预处理均发生在识别这些关键字之前。下表列出了标识C定义的关键字，其中加“ \* ”号的均是C99标准中新增加的。

**标准C的关键字**

auto \_Bool\* break case char \_Complex\*

const continue default restrict\* do double

else enum extern float for goto

if \_Imaginary\* inline\* int long register

return short signed sizeof static struct

switch typedef union unsigned void volatile

while

此外，有些C编译系统中还规定了其他的关键字。在不同的系统中，附加的关键字是不同的。例如，Borland C中附加的关键字有asm，cdecl，far，huge，interrupt，near，pascal，其含义可参考具体编译程序的有关资料。

**5. 分隔符**

分隔符统称为空白字符，包括空格符、制表符、换行符、换页符及注释符，它们在语法上仅起分隔单词的作用。当程序中两个相邻的单词之间如果不用分隔符就不能将两者区分开时，则必须加分隔符（通常用空格符）。为了使程序更加清晰，便于阅读，任何单词之间都可以加入适当的空格符。