C++ std::string 类型本质是什么？什么是字符串字面量？

C++ 字符串 string

C++ std::string 类型本质是什么？什么是字符串字面量？

问一下各位大佬：

1.std::string类型的变量学校老师说本质是char类型的数组，然而班上一些大佬否认了这一观点，并且我在测试的时候也发现point(char* a)构造函数无法传入std::string类型变量。

请问std::string变量真的本质是char类型数组吗？

2.就point(const char* a)构造函数可以直接传入字符串字面量：point a("safsf")而不能传入：std::string c="asdfs" point a(c)问题
想问一下：字符串字面量与字符串的区别？

关注问题 写回答 邀请回答 好问题 3 2 条评论 分享 ··· 收起

14 个回答 默认排序

exiledkingcc
专业修bug

d41d8c 等 10 人赞同了该回答

不要纠结什么“本质”，因为本质这种模糊的概念对于理解并没有什么帮助。

C++中的std::string内部就是char数组。但是本质是什么？不要去纠结了。反正归根到底，都是0和1。

至于point(const char*)可以传入“hello”但是不能传入std::string是因为前者的类型是匹配的，后者不是的。

字符串字面量，你可以暂且理解为源代码里面双引号中的内容。它的类型是什么取决于C++标准。

发布于 2021-04-18 16:50

赞同 10 2 条评论 分享 收藏 喜欢 ···

szouc
构建世界观

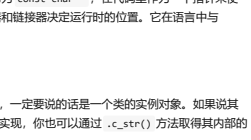
d41d8c 赞同了该回答

字面量顾名思义就是**人**看到字面就知道意思(信息)的量，字面量通常是人给出的计算机无法修改，所以又称为常量，它的符号也称为常量标识符，如 1，3.14，"hello"等。与此对应就是变量(对象)标识符，如 a，flag，str，**人**只看标识符符号是不能获取信息的。

无论是字面量还是变量，他们的符号都需要类型，否则人无法获取信息，如 "a" 是什么意思？类型是英文字母就是 ei，拼音就是 啊。标量的类型容易理解，但是复合量的类型就比较麻烦了，在 C语言中通常将复合量转换为标量处理，这个过程中不可避免的需要将复合量类型和标量类型的转换，如 C语言中的数组处理通常转换为指针处理，因此数组类型和指针类型有隐式转换。而在 c++ 中既保留了 c 的处理方式，也构建了自己的 class 处理方式，但这两种方式不兼容。

发布于 2021-04-18 19:49

赞同 1 添加评论 分享 收藏 喜欢 ···



rayhunter
腾讯游戏 游戏客户端开发工程师

13 人赞同了该回答

刚好看到了，这题我来试着答一下。一丝不漏的话完全理解清楚的话，C++水平那需要相当高，因为std::string其实挺复杂，毕竟标准库用用的最多容器之一。

1.先说你说的：***字符串字面量类型***，这个实际叫作字符串常量，比如**"hello"**，它的类型是**const char [6]**,而非**const char***类型，在模板传参要格外小心，因为模板退化是是可以退化为const char*。再进一步:**"aaaa"**,它的类型是**const char [5]，而**const char [6]与const char [5]就不是一个类型，这点要注意哦，只是退化时都变成const char*，这时又可以传参进去。

2.再说说：***字符串字面量编译后位置***，这玩意是常量，位于二进制的只读常量区，和虚函数在一个地方，通常反编译时，根据这些信息找突破口，一些对安全要求高的程序，都很谨慎使用常量字符串。

3.std::string 本质是个模板类，更进一步是**std::basic_string<char>** 的重定义，既然是个类，那么就可以有函数，可以有成员，有一些类型标识，有了这些包装，就可以有强大的功能，比如拼接，插入，查找，Hash等，单纯的C语言字符串做不到这些。不过恰巧这个模板还挺复杂，里面一迭代器，内存分配器，重载运算符，各种typedef/using，萃取。。。，我们简单点说，string本身很不大，成员不多，不同平台实现还有点区别，VS 2019只有一个成员是_Compressed_pair<_Alty, _Scary_val> _Mypair;，但字符串真正内存是堆区申请出来的，由成员_Mypair._Myval2._Myptr()指过去，加上模板类重载了操作符，一定程度可以const char*，以及char [N]进行转换。

新加VS2019下string的剖析

编辑于 2021-05-12 14:08

赞同 13 4 条评论 分享 收藏 喜欢 ···

nekosu
不务正业的退役Oier

4 人赞同了该回答

std::string（不是String）

这个是C++的字符串类型，在 `#include<string>` 中，是一个标准库定义实现的**类**。

字符串字面量

指通过双引号及其前后缀形成的**立即值**，类型通常为 `const char *` ，在代码里作为一个指针来使用。该地址处存放该字符串实际的内容，由编译器和链接器决定运行时的位置。它在语言中与 `1` ， `3.14` 的地位是类似的。

回到你的问题上来。

`string` 变量的本质是啥？这个问题没有什么意义，一定要说的话是一个类的实例对象。如果说其内部实现的话，一般是通过动态分配的 `char` 数组实现，你也可以通过 `.c_str()` 方法取得其内部的字符串指针。

之所以不能将 `string` 对象直接传给你的构造函数，是因为 `c++` 是一个强类型的语言，而 `string` 类并不是 `const char *` ，且不存在对其的隐式转换函数（类可以实现向其它类型的隐式转换，并且可以据此来匹配函数需求），因此不能将其作为参数传入。

确实，c++不算强类型，是静态弱类型语言。

编辑于 2021-04-18 21:40

赞同 4 1 条评论 分享 收藏 喜欢 ···

Xi Yang
被VS和MinGW轮流喂屎

d41d8c 等 6 人赞同了该回答

删除

编辑于 2021-12-31 13:22

赞同 6 添加评论 分享 收藏 喜欢 ···

Tony
Distributed Small Datum in Sky

zh.wikipedia.org/wiki/S...

发布于 2021-04-18 18:18

赞同 添加评论 分享 收藏 喜欢 ···

CyanCloverFern
我们坚信，仍有人为了理想而活着。

1 人赞同了该回答

1. 请问string变量真的本质是char类型数组吗？

不是，string不是pod的甚至不是trival（平凡）的，而char数组/const char数组即pod又平凡。

这种区别导致即使C++类型系统放宽也不能让你用string代替const char^Q*。

string在微软的典型构造，在基类中有一个机器字长的alloc用指针，value部分除开实际存储用的数据结构，还有两个长度为一个机器字长的size_t类型表示其它信息。

假如定义专用的类型转化？不行，下节表。

2. 想问一下：字符串字面量与字符串^Q的区别？

字符串是抽象概念，字符串字面量和string都是一种实现方式。他们直接没有定义隐式类型转换所以string和字符串字面量之间不能直接转换。

第一部分提及过value部分有一个实际存储用的数据结构，现在拿微软的典型实现看一看。

（手机我不会插代码块将就看。。。变量名瞎起的）

union {

char a[SIZE_MAX]

value_type aa[SIZE_MAX]

value_type* paa

}

SIZE_MAX取决于value_type长度，sizeof(value_type)大于16时取1，其他时候为16/sizeof(value_type)。

这里value_type可以送char_16t char_32t甚至other，并不能假定你使用的是8位char。

存储较大范围的value_type*是可变的，没有办法保证const char*正常使用（旧内存被析构，或者内容被修改）。

编辑于 2021-04-19 08:48

赞同 1 3 条评论 分享 收藏 喜欢 ··· 收起

小明
魔法，炼丹，搬砖

2 人赞同了该回答

c++标准（草案）没有规定std::string（实际上是std::basic_string<T>）具体如何实现。但是标准还是下了一个“定义”：std::basic_string<T>是存储并操作非数组平凡标准布局类型的char-like对象序列。

这么文绉绉的，简单来说，std::basic_string<T>维护的肯定不是一个数组，因为**数组是定长的**，std::basic_string<T>是可以改变存储的char-like数量。

也有答主提到了msvc的实现：std::basic_string<T>类内成员是：_Compressed_pair<_Alty, _Scary_val> _Mypair。再仔细发现，_Mypair封装的是对_Scary_val的操作，_Scary_val是_String_val<T>的特化。也就是说std::basic_string<T>封装的是_String_val<T>的操作。而_String_val<T>的成员有一个定长buf数组，和一个指针（还有一个alias数组做abi转换的）：于是实际上std::basic_string<T>分配空间的逻辑（本质）是：**对分配空间小的std::string用的是数组，对分配空间大的std::string用的是allocate，用指针管理。**

其实point(char* a)传参的疑惑，本质不是std::string类是什么的问题，而是隐式转换的问题：类类型一般不能隐式转换成基础类型（或者复合类型）。虽然C++不是算是强类型，不过不该转还是不给转。

编辑于 2021-04-23 15:05

赞同 2 添加评论 分享 收藏 喜欢 ··· 收起

shirayuki

1 人赞同了该回答

point(a.data())就行了

发布于 2021-04-18 19:46

赞同 1 添加评论 分享 收藏 喜欢 ··· 收起

godlike.hy
死程一枚

令人迷惑的根源来自c语言中，char类型数组同时用于保存字符串^Q和保存二进制数据，而char字符串表示前者，string是取后面的用法

在C++的大部分实现中，string字符串的内部都是一个vector，你可以认为是一个能自由变化长度的数组，而这个数组的成员类型可以是char，这就是你们老师说的“本质是char类型的数组”。对于存储ascii类型的string对象来说，调用方法c_str()或者data()方法可以返回这个数组存放数据的首地址。对于string来说，那个char型数组只是开了个同样大小的buffer而已

当然这个说法并不严谨，对于宽字符类型(wstring，是string类对象的unicode实现）其内部存储是vector<wchar_t>，也就不算是char型数组了

至于为啥printf()函数不能输入string，理由很简单，printf函数是c函数，他不支持c++参数。。。你试试C++版printf std::cout就能同时支持char型数组和string了,这是C++为了向下兼容C整出来的混乱之一。。。

至于字符串字面量和字符串的区别，raaay0608的回答应该是比较完善的了

跑个题，C++中由于向下兼容C，所以字符串这个名词每个人脑子里想的都不是同一个东西，所以最好的方法是问这个之前问清楚他说的字符串到底是char,wchar_t,string,UNICODE_STRING或者什么其他奇怪的玩意儿

由于C++实际有很高的自由度，想了解字符串字面量和字符串可以参考一下rust语言的str类型和string，这门语言显式区分了这两种数据。

发布于 2021-04-19 14:34

赞同 添加评论 分享 收藏 喜欢 ··· 收起

巴巴托斯
但风向是会转变的。

string的本质是char[] 指的是**可能的**内部用于**存放数据**的**类型**

发布于 2021-04-20 09:37

赞同 添加评论 分享 收藏 喜欢 ··· 收起

啊白菜
人极度不友善，发言极度不友好，对知乎社区极无贡献

1 人赞同了该回答

去看string的class定义就好了

肯定是一个标签类，指向一个真实数据块，数据块有字符串实际长度，保留空间，数据长度，引用计数这些信息

而实际上的赋值，则是两个string对象指向同一个数据块，然后引用计数为2

假设两个字符串拼接，赋值到其中一个字符串^Q，则应该产生一个新的数据块对象，让被赋值的字符串指向它，而被赋值的字符串原来指向的那个数据块引用计数-1

可以参考string class的operator =

在各个编译器实现也是不同的

发布于 2021-04-18 17:42

赞同 1 添加评论 分享 收藏 喜欢 ··· 收起

C十十20年
华中科技大学 工学博士

char *类型的字面量："abc"。

string类型的字面量："abc"s。

如果编译器必须识别一个字面值的类型，就说明这个类型就像int类型一样，快要成为标准类型了。但int类型已经进入保留字表，而string类型^Q还没进入。类似的准标准类型还有string_view以及initialize_list，他们的字面量分别有"abc"sv和{1,2,3,4}。

编辑于 2021-04-19 09:04

赞同 添加评论 分享 收藏 喜欢 ··· 收起

茄子
El Psy Congroo

你老师是对的。他比你们看的更深一步。

但是你也知道是本质相等，而语法层面不等。

struct A{char *b}; A a;

void print(const char *a);

虽然A的本质是一个char *b，但是你依然无法直接这么写print(a)；

发布于 2021-04-19 18:41

赞同 添加评论 分享 收藏 喜欢 ··· 收起

C++ std::string 类型本质原是什么？什么是字符串字面量？