

---

# Extending and Embedding Python

发布 3.8.0b1

Guido van Rossum  
and the Python development team

七月 04, 2019

Python Software Foundation  
Email: [docs@python.org](mailto:docs@python.org)



<b>1</b>	<b>推荐的第三方工具</b>	<b>3</b>
<b>2</b>	<b>不使用第三方工具创建扩展</b>	<b>5</b>
2.1	使用 C 或 C++ 扩展 Python . . . . .	5
2.2	自定义扩展类型：教程 . . . . .	22
2.3	定义扩展类型：已分类主题 . . . . .	45
2.4	构建 C/C++ 扩展 . . . . .	55
2.5	在 Windows 平台编译 C 和 C++ 扩展 . . . . .	57
<b>3</b>	<b>在更大的应用程序中嵌入 CPython 运行时</b>	<b>61</b>
3.1	Embedding Python in Another Application . . . . .	61
<b>A</b>	<b>术语对照表</b>	<b>67</b>
<b>B</b>	<b>文档说明</b>	<b>79</b>
B.1	Python 文档贡献者 . . . . .	79
<b>C</b>	<b>历史和许可证</b>	<b>81</b>
C.1	软件历史 . . . . .	81
C.2	访问 Python 或以其他方式使用 Python 的条款和条件 . . . . .	82
C.3	集成软件的许可和认可 . . . . .	85
<b>D</b>	<b>版权</b>	<b>99</b>
	<b>索引</b>	<b>101</b>



本文档描述了如何使用 C 或 C++ 编写模块以使用新模块来扩展 Python 解释器的功能。这些模块不仅可以定义新的函数，还可以定义新的对象类型及其方法。该文档还描述了如何将 Python 解释器嵌入到另一个应用程序中，以用作扩展语言。最后，它展示了如何编译和链接扩展模块，以便它们可以动态地（在运行时）加载到解释器中，如果底层操作系统支持此特性的话。

本文档假设你具备有关 Python 的基本知识。有关该语言的非正式介绍，请参阅 [tutorial-index](#)。[reference-index](#) 给出了更正式的语言定义。[library-index](#) 包含现有的对象类型、函数和模块（内置和用 Python 编写）的文档，使语言具有广泛的应用范围。

关于整个 Python/C API 的详细介绍，请参阅独立的 [c-api-index](#)。



# CHAPTER 1

---

## 推荐的第三方工具

---

本指南仅介绍了作为此 CPython 版本的一部分提供的创建扩展的基本工具。第三方工具，如 [Cython](#)、[cffi](#)、[SWIG](#) 和 [Numba](#) 提供了更简单和更复杂的方法来为 Python 创建 C 和 C++ 扩展。

**参见：**

**Python Packaging User Guide: Binary Extensions** “Python Packaging User Guide” 不仅涵盖了几个简化二进制扩展创建的可用工具，还讨论了为什么首先创建扩展模块的各种原因。





---

## 不使用第三方工具创建扩展

---

本指南的这一部分包括在没有第三方工具帮助的情况下创建 C 和 C++ 扩展。它主要用于这些工具的创建者，而不是建议你创建自己的 C 扩展的方法。

### 2.1 使用 C 或 C++ 扩展 Python

如果你会用 C，添加新的 Python 内置模块会很简单。以下两件不能用 Python 直接做的事，可以通过 *extension modules* 来实现：实现新的内置对象类型；调用 C 的库函数和系统调用。

为了支持扩展，Python API（应用程序编程接口）定义了一系列函数、宏和变量，可以访问 Python 运行时系统的大部分内容。Python 的 API 可以通过在一个 C 源文件中引用 "Python.h" 头文件来使用。

扩展模块的编写方式取决与你的目的以及系统设置；下面章节会详细介绍。

---

**注解：** C 扩展接口特指 CPython，扩展模块无法在其他 Python 实现上工作。在大多数情况下，应该避免写 C 扩展，来保持可移植性。举个例子，如果你的用例调用了 C 库或系统调用，你应该考虑使用 `ctypes` 模块或 `cffi` 库，而不是自己写 C 代码。这些模块允许你写 Python 代码来接口 C 代码，而且可移植性更好。不知为何编译失败了。

---

#### 2.1.1 一个简单的例子

让我们创建一个扩展模块 `spam` (Monty Python 粉丝最喜欢的食物...) 并且想要创建对应 C 库函数 `system()`<sup>1</sup> 的 Python 接口。这个函数接受一个以 `null` 结尾的字符串参数并返回一个整数。我们希望可以在 Python 中如下方式调用此函数：

```
>>> import spam
>>> status = spam.system("ls -l")
```

---

<sup>1</sup> 这个函数的接口已经在标准模块 `os` 里了，这里作为一个简单而直接的例子。

首先创建一个 `spammodule.c` 文件。(传统上, 如果一个模块叫 `spam`, 则对应实现它的 C 文件叫 `spammodule.c`; 如果这个模块名字非常长, 比如 `spammify`, 则这个模块的文件可以直接叫 `spammify.c`。)

文件中开始的两行是:

```
#define PY_SSIZE_T_CLEAN
#include <Python.h>
```

这会导入 Python API (如果你喜欢, 你可以在这里添加描述模块目标和版权信息的注释)。

**注解:** 由于 Python 可能会定义一些影响某些系统上标准头文件的预处理器定义, 因此在包含任何标准头文件之前, 您 \* 必须 \* `include` 这个文件: `Python.h`。

推荐总是在 `Python.h` 前定义 `PY_SSIZE_T_CLEAN`。查看[提取扩展函数的参数](#)来了解这个宏的更多内容。

所有用户可见的符号都定义自 `Python.h` 中, 并拥有前缀 `Py` 或 `PY`, 除了那些已经定义在标准头文件的。为了方便, 以及由于其在 Python 解释器中广泛应用, "`Python.h`" 也包含了少量标准头文件: `<stdio.h>`, `<string.h>`, `<errno.h>` 和 `<stdlib.h>`。如果后面的头文件在你的系统上不存在, 还会直接声明函数 `malloc()`, `free()` 和 `realloc()`。

下面要做的事是将 C 函数添加到我们的扩展模块, 当 Python 表达式 `spam.system(string)` 被求值时函数将被调用 (我们很快就会看到它最终是如何被调用的):

```
static PyObject *
spam_system(PyObject *self, PyObject *args)
{
    const char *command;
    int sts;

    if (!PyArg_ParseTuple(args, "s", &command))
        return NULL;
    sts = system(command);
    return PyLong_FromLong(sts);
}
```

有个直接翻译参数列表的方法 (例如单独的 "`ls-l`") 到要传递给 C 函数的参数。C 函数总是有两个参数, 通常名字是 `self` 和 `args`。

对模块级函数, `self` 参数指向模块对象; 对于对象实例则指向方法。

`args` 参数是指向一个 Python 的 `tuple` 对象的指针, 其中包含参数。每个 `tuple` 项对应一个调用参数。这些参数也全都是 Python 对象 — 要在我们的 C 函数中使用它们就需要先将其转换为 C 值。Python API 中的函数 `PyArg_ParseTuple()` 会检查参数类型并将其转换为 C 值。它使用模板字符串确定需要的参数类型以及存储被转换的值的 C 变量类型。细节将稍后说明。

`PyArg_ParseTuple()` 正常返回非零, 并已经按照提供的地址存入了各个变量值。如果出错 (零) 则应该让函数返回 `NULL` 以通知解释器出错 (有如例子中看到的)。

### 2.1.2 关于错误和异常

一个 Python 解释器的常见惯例是, 函数发生错误时, 应该设置一个异常环境并返回错误值 (通常是 `NULL` 指针)。异常存储在解释器静态全局变量中, 如果为 `NULL`, 则没有发生异常。异常的第一个参数也需要保存在静态全局变量中, 也就是 `raise` 的第二个参数 (第二个参数到 `raise`)。第三个变量包含栈回溯信息。这三个变量等同于 Python 变量 `sys.exc_info()` (查看 Python 库参考的模块 `sys` 的章节)。这对于理解到底发生了什么错误是很重要的。

Python API 中定义了一些函数来设置这些变量。

最常用的就是 `PyErr_SetString()`。其参数是异常对象和 C 字符串。异常对象一般是像 `PyExc_ZeroDivisionError` 这样的预定义对象。C 字符串指明异常原因，并被转换为一个 Python 字符串对象存储为异常的“关联值”。

另一个有用的函数是 `PyErr_SetFromErrno()`，仅接受一个异常对象，异常描述包含在全局变量 `errno` 中。最通用的函数还是 `PyErr_SetObject()`，包含两个参数，分别为异常对象和异常描述。你不需要使用 `Py_INCREF()` 来增加传递到其他函数的参数对象的引用计数。

你可以通过 `PyErr_Occurred()` 获知当前异常，返回当前异常对象，如果确实没有则为 `NULL`。一般来说，你在调用函数时不需要调用 `PyErr_Occurred()` 检查是否发生了异常，你可以直接检查返回值。

当函数 *f* 调用另一个函数 *g* 时检测到后者出错了，*f* 自身将返回一个错误值（通常为 `NULL` 或 `-1`）。它不应调用某个 `PyErr_*` 函数——这种函数已经由 *g* 调用过了。然后 *f* 的调用者也应该返回一个错误提示它的调用者，同样不应调用 `PyErr_*`，依此类推——错误的最详细原因已经由首先检测到它的函数报告了。一旦这个错误到达了 Python 解释器的主循环，它将中断当前执行的 Python 代码并尝试找到由 Python 程序员所指定的异常处理。

（在某些情况下，当模块确实能够通过调用其它 `PyErr_*` 函数给出更加详细的错误消息，并且在这些情况是可以这样做的。但是按照一般规则，这是不必要的，并可能导致有关错误原因的信息丢失：大多数操作会由于种种原因而失败。）

想要忽略由一个失败的函数调用所设置的异常，异常条件必须通过调用 `PyErr_Clear()` 显式地被清除。C 代码应当调用 `PyErr_Clear()` 的唯一情况是如果它不想将错误传给解释器而是想完全由自己来处理它（可能是尝试其他方法，或是假装没有出错）。

每次失败的 `malloc()` 调用必须转换为一个异常。`malloc()`（或 `realloc()`）的直接调用者必须调用 `PyErr_NoMemory()` 来返回错误来提示。所有对象创建函数（例如 `PyLong_FromLong()`）已经这么做了，所以这个提示仅用于直接调用 `malloc()` 的情况。

还要注意的，除了 `PyArg_ParseTuple()` 等重要的例外，返回整数状态码的函数通常都是返回正值或零来表示成功，而以 `-1` 表示失败，如同 Unix 系统调用一样。

最后，当你返回一个错误指示器时要注意清理垃圾（通过为你已经创建的对象执行 `Py_XDECREF()` 或 `Py_DECREF()` 调用）！

选择引发哪个异常完全取决于你的喜好。所有内置的 Python 异常都有对应的预声明 C 对象，例如 `PyExc_ZeroDivisionError`，你可以直接使用它们。当然，你应当明智地选择异常——不要使用 `PyExc_TypeError` 来表示一个文件无法被打开（那大概应该用 `PyExc_IOError`）。如果参数列表有问题，`PyArg_ParseTuple()` 函数通常会引发 `PyExc_TypeError`。如果你想要一个参数的值必须处于特定范围之内或必须满足其他条件，则适宜使用 `PyExc_ValueError`。

你也可以为你的模块定义一个唯一的新异常。需要在文件前部声明一个静态对象变量，如：

```
static PyObject *SpamError;
```

以及初始化你的模块的初始化函数 (`PyInit_spam()`) 包含一个异常对象（先不管错误检查）：

```
PyMODINIT_FUNC
PyInit_spam(void)
{
    PyObject *m;

    m = PyModule_Create(&spammodule);
    if (m == NULL)
        return NULL;

    SpamError = PyErr_NewException("spam.error", NULL, NULL);
    Py_INCREF(SpamError);
    PyModule_AddObject(m, "error", SpamError);
}
```

（下页继续）

(续上页)

```

    return m;
}

```

注意实际的 Python 异常名字是 `spam.error`。`PyErr_NewException()` 函数使用 `Exception` 为基类创建一个类 (除非是使用另外一个类替代 `NULL`)。描述参考 `bltin-exceptions`。

同样注意的是创建类保存了 `SpamError` 的一个引用, 这是有意的。为了防止被垃圾回收掉, 否则 `SpamError` 随时会成为野指针。

一会讨论 `PyMODINIT_FUNC` 作为函数返回类型的用法。

`spam.error` 异常可以在扩展模块中抛出, 通过 `PyErr_SetString()` 函数调用, 如下:

```

static PyObject *
spam_system(PyObject *self, PyObject *args)
{
    const char *command;
    int sts;

    if (!PyArg_ParseTuple(args, "s", &command))
        return NULL;
    sts = system(command);
    if (sts < 0) {
        PyErr_SetString(SpamError, "System command failed");
        return NULL;
    }
    return PyLong_FromLong(sts);
}

```

### 2.1.3 回到例子

回到前面的例子, 你应该明白下面的代码:

```

if (!PyArg_ParseTuple(args, "s", &command))
    return NULL;

```

如果在参数列表中检测到错误, 将会返回 `NULL` (返回对象指针的函数的错误指示器), 依据 `PyArg_ParseTuple()` 所设置的异常。在其他情况下参数的字符串值会被拷贝到局部变量 `command`。这是一个指针赋值, 你不应该修改它所指向的字符串 (所以在标准 C 中, 变量 `command` 应当被正确地声明为 `const char *command`)。

下一个语句使用 UNIX 系统函数 `system()`, 传递给他的参数是刚才从 `PyArg_ParseTuple()` 取出的:

```

sts = system(command);

```

我们的 `spam.system()` 函数必须返回 `sts` 的值作为 Python 对象。这通过使用函数 `PyLong_FromLong()` 来实现。

```

return PyLong_FromLong(sts);

```

在这种情况下, 会返回一个整数对象, (这个对象会在 Python 堆里面管理)。

如果你的 C 函数没有有用的返回值 (返回 `void` 的函数), 则必须返回 `None`。(你可以用 `Py_RETURN_NONE` 宏来完成):

```
Py_INCREF(Py_None);
return Py_None;
```

`Py_None` 是一个 C 名字指定 Python 对象 `None`。这是一个真正的 PY 对象，而不是 `NULL` 指针。

### 2.1.4 模块方法表和初始化函数

为了展示 `spam_system()` 如何被 Python 程序调用。把函数声明为可以被 Python 调用，需要先定义一个方法表“method table”。

```
static PyMethodDef SpamMethods[] = {
    ...
    {"system", spam_system, METH_VARARGS,
     "Execute a shell command."},
    ...
    {NULL, NULL, 0, NULL} /* Sentinel */
};
```

注意第三个参数 (`METH_VARARGS`)，这个标志指定会使用 C 的调用惯例。可选值有 `METH_VARARGS`、`METH_VARARGS | METH_KEYWORDS`。值 0 代表使用 `PyArg_ParseTuple()` 的陈旧变量。

如果单独使用 `METH_VARARGS`，函数会等待 Python 传来 tuple 格式的的参数，并最终使用 `PyArg_ParseTuple()` 进行解析。

`METH_KEYWORDS` 值表示接受关键字参数。这种情况下 C 函数需要接受第三个 `PyObject *` 对象，表示字典参数，使用 `PyArg_ParseTupleAndKeywords()` 来解析出参数。

这个方法表必须被模块定义结构所引用。

```
static struct PyModuleDef spammodule = {
    PyModuleDef_HEAD_INIT,
    "spam", /* name of module */
    spam_doc, /* module documentation, may be NULL */
    -1, /* size of per-interpreter state of the module,
        or -1 if the module keeps state in global variables. */
    SpamMethods
};
```

这个结构体必须传递给解释器的模块初始化函数。初始化函数必须命名为 `PyInit_name()`，其中 *name* 是模块的名字，并应该定义为非 `static`，且在模块文件里：

```
PyMODINIT_FUNC
PyInit_spam(void)
{
    return PyModule_Create(&spammodule);
}
```

注意 `PyMODINIT_FUNC` 声明了函数作为 `PyObject *` 返回类型，声明任何平台的链接生命，以及给 C++ 生命函数的 `extern "C"`。

当 Python 程序首次导入模块 `spam` 时，`PyInit_spam()` 被调用。(查看后续注释了解嵌入 Python)。他会调用 `PyModule_Create()`，会返回模块对象，并插入内置函数对象到新创建的模块里，基于表 (一个 `PyMethodDef` 结构体的数组类型) 到模块定义。`PyModule_Create()` 返回一个指向刚创建模块的指针。也可能因为严重错误而中止，或返回 `NULL` 在模块无法初始化成功时。初始化函数必须返回模块对象给调用者，所以之后会被插入 `sys.modules`。

当嵌入 Python 时, `PyInit_spam()` 函数不会被自动调用, 除非放在 `PyImport_Inittab` 表里。要添加模块到初始化表, 使用 `PyImport_AppendInittab()`, 可选的跟着一个模块的导入。

```
int
main(int argc, char *argv[])
{
    wchar_t *program = Py_DecodeLocale(argv[0], NULL);
    if (program == NULL) {
        fprintf(stderr, "Fatal error: cannot decode argv[0]\n");
        exit(1);
    }

    /* Add a built-in module, before Py_Initialize */
    PyImport_AppendInittab("spam", PyInit_spam);

    /* Pass argv[0] to the Python interpreter */
    Py_SetProgramName(program);

    /* Initialize the Python interpreter.  Required. */
    Py_Initialize();

    /* Optionally import the module; alternatively,
       import can be deferred until the embedded script
       imports it. */
    PyImport_ImportModule("spam");

    ...

    PyMem_RawFree(program);
    return 0;
}
```

**注解:** 要从 `sys.modules` 删除实体或导入已编译模块到一个进程里的多个解释器 (或使用 `fork()` 而没用 `exec()`) 会在一些扩展模块上产生错误。扩展模块作者可以在初始化内部数据结构时给出警告。

更多关于模块的现实的例子包含在 Python 源码包的 `Modules/xxmodule.c` 中。这些文件可以用作你的代码模板, 或者学习。脚本 `modulator.py` 包含在源码发行版或 Windows 安装中, 提供了一个简单的 GUI, 用来声明需要实现的函数和对象, 并且可以生成供填入的模板。脚本在 `Tools/modulator/` 目录。查看 `README` 以了解用法。

**注解:** 不像我们的 `spam` 例子, `xxmodule` 使用了多阶段初始化 (Python3.5 开始引入), `PyInit_spam` 会返回一个 `PyModuleDef` 结构体, 然后创建的模块放到导入机制。细节参考 [PEP 489](#) 的多阶段初始化。

### 2.1.5 编译和链接

在你使用你的新写的扩展之前, 你还需要做两件事情: 使用 Python 系统来编译和链接。如果你使用动态加载, 这取决于你使用的操作系统的动态加载机制; 更多信息请参考编译扩展模块的章节 ([构建 C/C++ 扩展章节](#)), 以及在 Windows 上编译需要的额外信息 ([在 Windows 平台编译 C 和 C++ 扩展章节](#))。

如果你不使用动态加载, 或者想要让模块永久性的作为 Python 解释器的一部分, 就必须修改配置设置, 并重新构建解释器。幸运的是在 Unix 上很简单, 只需要把你的文件 (`spammodule.c` 为例) 放在解压缩源码发行包的 `Modules/` 目录下, 添加一行到 `Modules/Setup.local` 来描述你的文件:



```
spam spammodule.o
```

然后在顶层目录运行 **make** 来重新构建解释器。你也可以在 Modules/ 子目录使用 **make**，但是你必须先重建 Makefile 文件，然后运行 **make Makefile** 命令。（你每次修改 Setup 文件都需要这样操作。）

如果你的模块需要额外的链接，这些内容可以列出在配置文件里，举个实例：

```
spam spammodule.o -lX11
```

## 2.1.6 在 C 中调用 Python 函数

迄今为止，我们一直把注意力集中于让 Python 调用 C 函数，其实反过来也很有用，就是用 C 调用 Python 函数。这在回调函数中尤其有用。如果一个 C 接口使用回调，那么就要实现这个回调机制。

幸运的是，Python 解释器是比较方便回调的，并给标准 Python 函数提供了标准接口。（这里就不再详述解析 Python 代码作为输入的方式，如果有兴趣可以参考 Python/pythonmain.c 中的 -c 命令代码。）

调用 Python 函数，首先 Python 程序要传递 Python 函数对象。应该提供个函数（或其他接口）来实现。当调用这个函数时，用全局变量保存 Python 函数对象的指针，还要调用 (Py\_INCREF()) 来增加引用计数，当然不用全局变量也没什么关系。例如如下：

```
static PyObject *my_callback = NULL;

static PyObject *
my_set_callback(PyObject *dummy, PyObject *args)
{
    PyObject *result = NULL;
    PyObject *temp;

    if (PyArg_ParseTuple(args, "O:set_callback", &temp)) {
        if (!PyCallable_Check(temp)) {
            PyErr_SetString(PyExc_TypeError, "parameter must be callable");
            return NULL;
        }
        Py_XINCREF(temp);           /* Add a reference to new callback */
        Py_XDECREF(my_callback);   /* Dispose of previous callback */
        my_callback = temp;        /* Remember new callback */
        /* Boilerplate to return "None" */
        Py_INCREF(Py_None);
        result = Py_None;
    }
    return result;
}
```

这个函数必须使用 METH\_VARARGS 标志注册到解释器，这在模块方法表和初始化函数 章节会描述。PyArg\_ParseTuple() 函数及其参数的文档在提取扩展函数的参数。

Py\_XINCREF() 和 Py\_XDECREF() 这两个宏可以用来增加或减少对象的引用计数，即使参数是 NULL 指针，操作也是安全的（但在这个例子中 temp 永远不会为 NULL）。更多内容请参考引用计数 段落。

PyEval\_CallObject() 返回一个 Python 对象指针表示返回值。该函数有 2 个参数，都是指向 Python 对象的指针：Python 函数，和参数列表。参数列表必须是 tuple 对象，其长度是参数数量。要调用无参数的 Python 函数，可以传递 NULL 或空元组。要用唯一参数调用，传递单一元组。Py\_BuildValue() 返回元组，当其格式为字符串或多个编码时，例如：

```
int arg;
PyObject *arglist;
PyObject *result;
...
arg = 123;
...
/* Time to call the callback */
arglist = Py_BuildValue("(i)", arg);
result = PyObject_CallObject(my_callback, arglist);
Py_DECREF(arglist);
```

`PyObject_CallObject()` 返回 Python 对象指针，这也是 Python 函数的返回值。`PyObject_CallObject()` 是一个对其参数“引用计数无关”的函数。例子中新的元组创建用于参数列表，并且在 `PyObject_CallObject()` 之后立即使用了 `Py_DECREF()`。

`PyEval_CallObject()` 的返回值总是“新”的：要么是一个新建的对象；要么是已有对象，但增加了引用计数。所以除非你想把结果保存在全局变量中，你需要对这个值使用 `Py_DECREF()`，即使你对里面的内容（特别！）不感兴趣。

在你这么做之前，需要先检查返回值是否是 `NULL`。如果是，Python 函数会终止并抛出异常。如果 C 代码调用了从 Python 传入的函数 `PyObject_CallObject()`，因该立即返回错误来告知 Python 调用者，然后解释器会打印栈回溯，或者调用 Python 代码来处理这个异常。如果无法处理，异常会被 `PyErr_Clear()` 清除，例如：

```
if (result == NULL)
    return NULL; /* Pass error back */
...use result...
Py_DECREF(result);
```

依赖于具体的回调函数，你还要提供一个参数列表到 `PyEval_CallObject()`。在某些情况下参数列表是由 Python 程序提供的，通过接口再传到回调函数。这样就可以不改变形式直接传递。另外一些时候你要构造一个新的 tuple 来传递参数。最简单的方法就是 `Py_BuildValue()` 函数构造 tuple。例如，你要传递一个事件对象时可以用：

```
PyObject *arglist;
...
arglist = Py_BuildValue("(l)", eventcode);
result = PyObject_CallObject(my_callback, arglist);
Py_DECREF(arglist);
if (result == NULL)
    return NULL; /* Pass error back */
/* Here maybe use the result */
Py_DECREF(result);
```

注意 `Py_DECREF(arglist)` 所在处会立即调用，在错误检查之前。当然还要注意一些常规的错误，比如 `Py_BuildValue()` 可能会遭遇内存不足等等。

你还需要注意，用关键字参数调用 `PyObject_Call()`，需要支持普通参数和关键字参数。有如如上例子中，我们使用 `Py_BuildValue()` 来构造字典。

```
PyObject *dict;
...
dict = Py_BuildValue("{s:i}", "name", val);
result = PyObject_Call(my_callback, NULL, dict);
Py_DECREF(dict);
if (result == NULL)
    return NULL; /* Pass error back */
```

(下页继续)



(续上页)

```
/* Here maybe use the result */
Py_DECREF(result);
```

## 2.1.7 提取扩展函数的参数

函数 `PyArg_ParseTuple()` 的声明如下:

```
int PyArg_ParseTuple(PyObject *arg, const char *format, ...);
```

参数 *arg* 必须是一个元组对象, 包含从 Python 传递给 C 函数的参数列表。 *format* 参数必须是一个格式字符串, 语法请参考 Python C/API 手册中的 `arg-parsing`。 剩余参数是各个变量的地址, 类型要与格式字符串对应。

注意 `PyArg_ParseTuple()` 会检测他需要的 Python 参数类型, 却无法检测传递给他的 C 变量地址, 如果这里出错了, 可能会在内存中随机写入东西, 小心。

注意任何由调用者提供的 Python 对象引用是 借来的引用; 不要递减它们的引用计数!

一些调用的例子:

```
#define PY_SSIZE_T_CLEAN /* Make "s#" use Py_ssize_t rather than int. */
#include <Python.h>
```

```
int ok;
int i, j;
long k, l;
const char *s;
Py_ssize_t size;

ok = PyArg_ParseTuple(args, ""); /* No arguments */
/* Python call: f() */
```

```
ok = PyArg_ParseTuple(args, "s", &s); /* A string */
/* Possible Python call: f('whoops!') */
```

```
ok = PyArg_ParseTuple(args, "lls", &k, &l, &s); /* Two longs and a string */
/* Possible Python call: f(1, 2, 'three') */
```

```
ok = PyArg_ParseTuple(args, "(ii)s#", &i, &j, &s, &size);
/* A pair of ints and a string, whose size is also returned */
/* Possible Python call: f((1, 2), 'three') */
```

```
{
    const char *file;
    const char *mode = "r";
    int bufsize = 0;
    ok = PyArg_ParseTuple(args, "s|si", &file, &mode, &bufsize);
    /* A string, and optionally another string and an integer */
    /* Possible Python calls:
       f('spam')
       f('spam', 'w')
       f('spam', 'wb', 100000) */
}
```

```
{
    int left, top, right, bottom, h, v;
    ok = PyArg_ParseTuple(args, "((ii)(ii))(ii)",
        &left, &top, &right, &bottom, &h, &v);
    /* A rectangle and a point */
    /* Possible Python call:
       f(((0, 0), (400, 300)), (10, 10)) */
}
```

```
{
    Py_complex c;
    ok = PyArg_ParseTuple(args, "D:myfunction", &c);
    /* a complex, also providing a function name for errors */
    /* Possible Python call: myfunction(1+2j) */
}
```

## 2.1.8 给扩展函数的关键字参数

函数 `PyArg_ParseTupleAndKeywords()` 声明如下:

```
int PyArg_ParseTupleAndKeywords(PyObject *arg, PyObject *kwdict,
                                const char *format, char *kwlist[], ...);
```

参数 `arg` 和 `format` 定义同 `PyArg_ParseTuple()`。参数 `kwdict` 是关键字字典, 用于接受运行时传来的关键字参数。参数 `kwlist` 是一个 `NULL` 结尾的字符串, 定义了可以接受的参数名, 并从左到右与 `format` 中各个变量对应。如果执行成功 `PyArg_ParseTupleAndKeywords()` 会返回 `true`, 否则返回 `false` 并抛出异常。

**注解:** 嵌套的元组在使用关键字参数时无法生效, 不在 `kwlist` 中的关键字参数会导致 `TypeError` 异常。

如下是使用关键字参数的例子模块, 作者是 Geoff Philbrick ([phibrick@hks.com](mailto:phibrick@hks.com)):

```
#define PY_SSIZE_T_CLEAN /* Make "s#" use Py_ssize_t rather than int. */
#include <Python.h>

static PyObject *
keywdarg_parrot(PyObject *self, PyObject *args, PyObject *keywds)
{
    int voltage;
    const char *state = "a stiff";
    const char *action = "vroom";
    const char *type = "Norwegian Blue";

    static char *kwlist[] = {"voltage", "state", "action", "type", NULL};

    if (!PyArg_ParseTupleAndKeywords(args, keywds, "i|sss", kwlist,
        &voltage, &state, &action, &type))
        return NULL;

    printf("-- This parrot wouldn't %s if you put %i Volts through it.\n",
        action, voltage);
    printf("-- Lovely plumage, the %s -- It's %s!\n", type, state);

    Py_RETURN_NONE;
}
```

(下页继续)

(续上页)

```

}

static PyMethodDef keywdarg_methods[] = {
    /* The cast of the function is necessary since PyCFunction values
     * only take two PyObject* parameters, and keywdarg_parrot() takes
     * three.
     */
    {"parrot", (PyCFunction)(void*)(void)keywdarg_parrot, METH_VARARGS | METH_
↪KEYWORDS,
    "Print a lovely skit to standard output."},
    {NULL, NULL, 0, NULL} /* sentinel */
};

static struct PyModuleDef keywdargmodule = {
    PyModuleDef_HEAD_INIT,
    "keywdarg",
    NULL,
    -1,
    keywdarg_methods
};

PyMODINIT_FUNC
PyInit_keywdarg(void)
{
    return PyModule_Create(&keywdargmodule);
}

```

## 2.1.9 构造任意值

这个函数与 `PyArg_ParseTuple()` 很相似，声明如下：

```
PyObject *Py_BuildValue(const char *format, ...);
```

接受一个格式字符串，与 `PyArg_ParseTuple()` 相同，但是参数必须是原变量的地址指针（输入给函数，而非输出）。最终返回一个 Python 对象适合于返回 C 函数调用给 Python 代码。

一个与 `PyArg_ParseTuple()` 的不同是，后面可能需要的要求返回一个元组（Python 参数里该包总是在内部描述为元组），比如用于传递给其他 Python 函数以参数。`Py_BuildValue()` 并不总是生成元组，在多于 1 个参数时会生成元组，而如果没有参数则返回 `None`，一个参数则直接返回该参数的对象。如果要求强制生成一个长度为空的元组，或包含一个元素的元组，需要在格式字符串中加上括号。

例子（左侧是调用，右侧是 Python 值结果）：

<code>Py_BuildValue("")</code>	<code>None</code>
<code>Py_BuildValue("i", 123)</code>	<code>123</code>
<code>Py_BuildValue("iii", 123, 456, 789)</code>	<code>(123, 456, 789)</code>
<code>Py_BuildValue("s", "hello")</code>	<code>'hello'</code>
<code>Py_BuildValue("y", "hello")</code>	<code>b'hello'</code>
<code>Py_BuildValue("ss", "hello", "world")</code>	<code>('hello', 'world')</code>
<code>Py_BuildValue("s#", "hello", 4)</code>	<code>'hell'</code>
<code>Py_BuildValue("y#", "hello", 4)</code>	<code>b'hell'</code>
<code>Py_BuildValue("()")</code>	<code>()</code>
<code>Py_BuildValue("(i)", 123)</code>	<code>(123,)</code>
<code>Py_BuildValue("(ii)", 123, 456)</code>	<code>(123, 456)</code>
<code>Py_BuildValue("(i,i)", 123, 456)</code>	<code>(123, 456)</code>

(下页继续)

(续上页)

Py_BuildValue("[i,i]", 123, 456)	[123, 456]
Py_BuildValue("{s:i,s:i}", "abc", 123, "def", 456)	{'abc': 123, 'def': 456}
Py_BuildValue("((ii)(ii)) (ii)", 1, 2, 3, 4, 5, 6)	(( (1, 2), (3, 4)), (5, 6))

## 2.1.10 引用计数

在 C/C++ 语言中，程序员负责动态分配和回收堆 (heap) 当中的内存。在 C 里，通过函数 `malloc()` 和 `free()` 来完成。在 C++ 里是操作 `new` 和 `delete` 来实现相同的功能。

每个由 `malloc()` 分配的内存块，最终都要由 `free()` 退回到可用内存池里面去。而调用 `free()` 的时机非常重要，如果一个内存块忘了 `free()` 则会导致内存泄漏，这块内存存在程序结束前将无法重新使用。这叫做内存泄漏。而如果对同一内存块 `free()` 了以后，另外一个指针再次访问，则再次使用 `malloc()` 复用这块内存会导致冲突。这叫做野指针。等同于使用未初始化的数据，core dump，错误结果，神秘的崩溃等。

内存泄露往往发生在一些并不常见的代码流程上面。比如一个函数申请了内存以后，做了些计算，然后释放内存块。现在一些对函数的修改可能增加对计算的测试并检测错误条件，然后过早的从函数返回了。这很容易忘记在退出前释放内存，特别是后期修改的代码。这种内存泄漏，一旦引入，通常很长时间都难以检测到，错误退出被调用的频度较低，而现代电脑又有非常巨大的虚拟内存，所以泄漏仅在长期运行或频繁调用泄漏函数时才会变得明显。因此，有必要避免内存泄漏，通过代码规范会策略来最小化此类错误。

Python 通过 `malloc()` 和 `free()` 包含大量的内存分配和释放，同样需要避免内存泄漏和野指针。他选择的方法就是引用计数。其原理比较简单：每个对象都包含一个计数器，计数器的增减与对象引用的增减直接相关，当引用计数为 0 时，表示对象已经没有存在的意义了，对象就可以删除了。

另一个叫法是自动垃圾回收。(有时引用计数也被看作是垃圾回收策略，于是这里的“自动”用以区分两者)。自动垃圾回收的优点是用户不需要明确的调用 `free()`。(另一个优点是改善速度或内存使用，然而这并不难)。缺点是对 C，没有可移植的自动垃圾回收器，而引用计数则可以可移植的实现(只要 `malloc()` 和 `free()` 函数是可用的，这也是 C 标准担保的)。也许以后有一天会出现可移植的自动垃圾回收器，但在此前我们必须与引用计数一起工作。

Python 使用传统的引用计数实现，也提供了循环监测器，用以检测引用循环。这使得应用无需担心直接或间接的创建了循环引用，这是引用计数垃圾收集的一个弱点。引用循环是对象(可能直接)的引用了本身，所以循环中的每个对象的引用计数都不是 0。典型的引用计数实现无法回收处于引用循环中的对象，或者被循环所引用的对象，哪怕没有循环以外的引用了。

循环探测器可以检测垃圾循环并回收。gc 模块提供了方法运行探测器 (`collect()` 函数)，而且可以在运行时配置禁用探测器。循环探测器被当作可选组件，默认是包含的，也可以在构建时禁用，在 Unix 平台(包括 Mac OS X)使用 `--without-cycle-gc` 选项到 **configure** 脚本。如果循环探测器被禁用，gc 模块就不可用了。

## Python 中的引用计数

有两个宏 `Py_INCREF(x)` 和 `Py_DECREF(x)`，会处理引用计数的增减。`Py_DECREF()` 也会在引用计数到达 0 时释放对象。为了灵活，并不会直接调用 `free()`，而是通过对象的类型对象的函数指针来调用。为了这个目的(或其他的)，每个对象同时包含一个指向自身类型对象的指针。

最大的问题依旧：何时使用 `Py_INCREF(x)` 和 `Py_DECREF(x)`？我们首先引入一些概念。没有人“拥有”一个对象，你可以拥有一个引用到一个对象。一个对象的引用计数定义为拥有引用的数量。引用的拥有者有责任调用 `Py_DECREF()`，在引用不再需要时。引用的拥有关系可以被传递。有三种办法来处置拥有的引用：传递、存储、调用 `Py_DECREF()`。忘记处置一个拥有的引用会导致内存泄漏。

还可以借用<sup>2</sup>一个对象的引用。借用的引用不应该调用 `Py_DECREF()`。借用者必须确保不能持有对象超过拥有者借出的时间。在拥有者处置对象后使用借用的引用是有风险的，应该完全避免<sup>3</sup>。

借用相对于引用的优点是你无需担心整条路径上代码的引用，或者说，通过借用你无需担心内存泄漏的风险。借用的缺点是一些看起来正确代码上的借用可能会在拥有者处置后使用对象。

借用可以变为拥有引用，通过调用 `Py_INCREF()`。这不会影响已经借出的拥有者的状态。这回创建一个新的拥有引用，并给予完全的拥有者责任（新的拥有者必须恰当的处置引用，就像之前的拥有者那样）。

## 拥有规则

当一个对象引用传递进出一个函数时，函数的接口应该指定拥有关系的传递是否包含引用。

大多数函数返回一个对象的引用，并传递引用拥有关系。通常，所有创建对象的函数，例如 `PyLong_FromLong()` 和 `Py_BuildValue()`，会传递拥有关系给接收者。即便是对象不是真正新的，你仍然可以获得对象的新引用。一个实例是 `PyLong_FromLong()` 维护了一个流行值的缓存，并可以返回已缓存项目的新引用。

很多另一个对象提取对象的函数，也会传递引用关系，例如 `PyObject_GetAttrString()`。这里的情况不够清晰，一些不太常用的例程是例外的 `PyTuple_GetItem()`，`PyList_GetItem()`，`PyDict_GetItem()`，`PyDict_GetItemString()` 都是返回从元组、列表、字典里借用的引用。

函数 `PyImport_AddModule()` 也会返回借用的引用，哪怕可能会返回创建的对象：这个可能因为一个拥有的引用对象是存储在 `sys.modules` 里。

当你传递一个对象引用到另一个函数时，通常函数是借用出去的。如果需要存储，就使用 `Py_INCREF()` 来变成独立的拥有者。这个规则有两个重要的例外：`PyTuple_SetItem()` 和 `PyList_SetItem()`。这些函数接受传递来的引用关系，哪怕会失败！（注意 `PyDict_SetItem()` 及其同类不会接受引用关系，他们是“正常的”）。

当一个 C 函数被 Python 调用时，会从调用方传来的参数借用引用。调用者拥有对象的引用，所以借用的引用生命周期可以保证到函数返回。只要当借用的引用需要存储或传递时，就必须转换为拥有的引用，通过调用 `Py_INCREF()`。

Python 调用从 C 函数返回的对象引用时必须是有拥有的引用——拥有关系被从函数传递给调用者。

## 危险的薄冰

有少数情况下，借用的引用看起来无害，但却可能导致问题。这通常是因为解释器的隐式调用，并可能导致引用拥有者处置这个引用。

首先需要特别注意的情况是使用 `Py_DECREF()` 到一个无关对象，而这个对象的引用是借用自一个列表的元素。举个实例：

```
void
bug(PyObject *list)
{
    PyObject *item = PyList_GetItem(list, 0);

    PyList_SetItem(list, 1, PyLong_FromLong(0L));
    PyObject_Print(item, stdout, 0); /* BUG! */
}
```

这个函数首先借用一个引用 `list[0]`，然后替换 `list[1]` 为值 0，最后打印借用的引用。看起来无害是吧，但却不是。

<sup>2</sup> 术语“借用”一个引用是不完全正确的：拥有者仍然有引用的拷贝。

<sup>3</sup> 检查引用计数至少为 1 没有用，引用计数本身可以在已经释放的内存里，并有可能被其他对象所用。

我们跟着控制流进入 `PyList_SetItem()`。列表拥有者引用了其所有成员，所以当成员 1 被替换时，就必须处置原来的成员 1。现在假设原来的成员 1 是用户定义类的实例，且假设这个类定义了 `__del__()` 方法。如果这个类实例的引用计数是 1，那么处置动作就会调用 `__del__()` 方法。

既然是 Python 写的，`__del__()` 方法可以执行任意 Python 代码。是否可能在 `bug()` 的 `item` 废止引用呢，是的。假设列表传递到 `bug()` 会被 `__del__()` 方法所访问，就可以执行一个语句来实现 `del list[0]`，然后假设这是最后一个对对象的引用，就需要释放内存，从而使得 `item` 无效化。

解决方法是，当你知道了问题的根源，就容易了：临时增加引用计数。正确版本的函数代码如下：

```
void
no_bug(PyObject *list)
{
    PyObject *item = PyList_GetItem(list, 0);

    Py_INCREF(item);
    PyList_SetItem(list, 1, PyLong_FromLong(0L));
    PyObject_Print(item, stdout, 0);
    Py_DECREF(item);
}
```

这是个真实的故事。一个旧版本的 Python 包含了这个 bug 的变种，而一些人花费了大量时间在 C 调试器上去寻找为什么 `__del__()` 方法会失败。

这个问题的第二种情况是借用的引用涉及线程的变种。通常，Python 解释器里多个线程无法进入对方的路径，因为有个全局锁保护着 Python 整个对象空间。但可以使用宏 `Py_BEGIN_ALLOW_THREADS` 来临时释放这个锁，重新获取锁用 `Py_END_ALLOW_THREADS`。这通常围绕在阻塞 I/O 调用外，使得其他线程可以在等待 I/O 期间使用处理器。显然，如下函数会跟之前那个有一样的问题：

```
void
bug(PyObject *list)
{
    PyObject *item = PyList_GetItem(list, 0);
    Py_BEGIN_ALLOW_THREADS
    ...some blocking I/O call...
    Py_END_ALLOW_THREADS
    PyObject_Print(item, stdout, 0); /* BUG! */
}
```

## NULL 指针

通常，函数接受对象引用作为参数，而非期待你传入 `NULL` 指针，你非这么干会导致 `dump core` (或者之后导致 `core dumps`)。函数返回对象引用时，返回的 `NULL` 用以指示发生了异常。`NULL` 参数的理由在从其他函数接收时并未测试，如果每个函数都测试 `NULL`，就会导致大量的冗余测试，并使得代码运行更慢。

好的方法是仅在“源头”测试 `NULL`，当一个指针可能是 `NULL` 时，例如 `malloc()` 或者从一个可能抛出异常的函数。

宏 `Py_INCREF()` 和 `Py_DECREF()` 不会检查 `NULL` 指针。但他们的变种 `Py_XINCREF()` 和 `Py_XDECREF()` 会检查。

用以检查对象类型的宏 (`Pytype_Check()`) 不会检查 `NULL` 指针，有很多代码会多次测试一个对象是否是预期的类型，这可能产生冗余的测试。而 `NULL` 检查没有冗余。

C 函数调用机制会确保传递到 C 函数的参数列表 (例如 `args`) 不会是 `NULL`，实际上会确保总是元组<sup>4</sup>。

把 `NULL` 指针转义给 Python 用户是个严重的错误。

<sup>4</sup> 当你使用“旧式”风格调用约定时，这些保证不成立，尽管这依旧存在于很多旧代码中。



### 2.1.11 在 C++ 中编写扩展

还可以在 C++ 中编写扩展模块，只是有些限制。如果主程序 (Python 解释器) 是使用 C 编译器来编译和链接的，全局或静态对象的构造器就不能使用。而如果是 C++ 编译器来链接的就没有这个问题。函数会被 Python 解释器调用 (通常就是模块初始化函数) 必须声明为 `extern "C"`。而是否在 `extern "C" {...}` 里包含 Python 头文件则不是那么重要，因为如果定义了符号 `__cplusplus` 则已经是这么声明的了 (所有现代 C++ 编译器都会定义这个符号)。

### 2.1.12 给扩展模块提供 C API

很多扩展模块提供了新的函数和类型供 Python 使用，但有时扩展模块里的代码也可以被其他扩展模块使用。例如，一个扩展模块可以实现一个类型“collection”看起来是没有顺序的。就像是 Python 列表类型，拥有 C API 允许扩展模块来创建和维护列表，这个新的集合类型可以有一堆 C 函数用于给其他扩展模块直接使用。

开始看起来很简单：只需要编写函数 (无需声明为 `static`)，提供恰当的头文件，以及 C API 的文档。实际上在所有扩展模块都是静态链接到 Python 解释器时也是可以正常工作的。当模块以共享库链接时，一个模块中的符号定义对另一个模块不可见。可见的细节依赖于操作系统，一些系统的 Python 解释器使用全局命名空间 (例如 Windows)，有些则在链接时需要一个严格的已导入符号列表 (一个例子是 AIX)，或者提供可选的不同策略 (如 Unix 系列)。即便是符号是全局可见的，你要调用的模块也可能尚未加载。

可移植性需要不能对符号可见性做任何假设。这意味着扩展模块里的所有符号都应该声明为 `static`，除了模块的初始化函数，来避免与其他扩展模块的命名冲突 (在段落 [模块方法表和初始化函数](#) 中讨论)。这意味着符号应该 必须通过其他导出方式来供其他扩展模块访问。

Python 提供了一个特别的机制来传递 C 级别信息 (指针)，从一个扩展模块到另一个：Capsules。一个 Capsule 是一个 Python 数据类型，会保存指针 (`void *`)。Capsule 只能通过其 C API 来创建和访问，但可以像其他 Python 对象一样的传递。通常，我们可以指定一个扩展模块命名空间的名字。其他扩展模块可以导入这个模块，获取这个名字的值，然后从 Capsule 获取指针。

Capsule 可以用多种方式导出 C API 给扩展模块。每个函数可以用自己的 Capsule，或者所有 C API 指针可以存储在一个数组里，数组地址再发布给 Capsule。存储和获取指针也可以用多种方式，供客户端模块使用。

使用的方法，对 Capsule 的名字很重要。函数 `PyCapsule_New()` 会接受一个名字参数 (`const char *`)，你可以传入 `NULL` 给名字，但强烈建议指定个名字。恰当的命名 Capsule 提供了一定程度的运行时类型安全；而却没有可行的方法来告知我们一个未命名的 Capsule。

通常来说，Capsule 用于暴露 C API，其名字应该遵循如下规范：

```
modulename.attributename
```

便利函数 `PyCapsule_Import()` 可以方便的载入通过 Capsule 提供的 C API，仅在 Capsule 的名字匹配时。这个行为为 C API 用户提供了高度的确定性来载入正确的 C API。

如下例子展示了将大部分负担交由导出模块作者的方法，适用于常用的库模块。其会存储所有 C API 指针 (例子里只有一个) 在 `void` 指针的数组里，并使其值变为 Capsule。对应的模块头文件提供了宏来管理导入模块和获取 C API 指针；客户端模块只需要在访问 C API 前调用这个宏即可。

导出的模块修改自 `spam` 模块，来自一个 [简单的例子](#) 段落。函数 `spam.system()` 不会直接调用 C 库函数 `system()`，但一个函数 `PySpam_System()` 会负责调用，当然现实中会更复杂些 (例如添加“spam”到每个命令)。函数 `PySpam_System()` 也会导出给其他扩展模块。

函数 `PySpam_System()` 是个纯 C 函数，声明 `static` 就像其他地方那样：

```
static int
PySpam_System(const char *command)
{
    return system(command);
}
```

函数 `spam_system()` 按照如下方式修改:

```
static PyObject *
spam_system(PyObject *self, PyObject *args)
{
    const char *command;
    int sts;

    if (!PyArg_ParseTuple(args, "s", &command))
        return NULL;
    sts = PySpam_System(command);
    return PyLong_FromLong(sts);
}
```

在模块开头, 在此行后:

```
#include <Python.h>
```

添加另外两行:

```
#define SPAM_MODULE
#include "spammodule.h"
```

`#define` 用于告知头文件需要包含给导出的模块, 而不是客户端模块。最终, 模块的初始化函数必须负责初始化 C API 指针数组:

```
PyMODINIT_FUNC
PyInit_spam(void)
{
    PyObject *m;
    static void *PySpam_API[PySpam_API_pointers];
    PyObject *c_api_object;

    m = PyModule_Create(&spammodule);
    if (m == NULL)
        return NULL;

    /* Initialize the C API pointer array */
    PySpam_API[PySpam_System_NUM] = (void *)PySpam_System;

    /* Create a Capsule containing the API pointer array's address */
    c_api_object = PyCapsule_New((void *)PySpam_API, "spam._C_API", NULL);

    if (c_api_object != NULL)
        PyModule_AddObject(m, "_C_API", c_api_object);
    return m;
}
```

注意 `PySpam_API` 声明为 `static`; 此外指针数组会在 `PyInit_spam()` 结束后消失!

头文件 `spammodule.h` 里的一堆工作, 看起来如下所示:

```
#ifndef Py_SPAMMODULE_H
#define Py_SPAMMODULE_H
#ifdef __cplusplus
extern "C" {
#endif
```

(下页继续)



(续上页)

```

/* Header file for spammodule */

/* C API functions */
#define PySpam_System_NUM 0
#define PySpam_System_RETURN int
#define PySpam_System_PROTO (const char *command)

/* Total number of C API pointers */
#define PySpam_API_pointers 1

#ifdef SPAM_MODULE
/* This section is used when compiling spammodule.c */

static PySpam_System_RETURN PySpam_System PySpam_System_PROTO;

#else
/* This section is used in modules that use spammodule's API */

static void **PySpam_API;

#define PySpam_System \
    (*(PySpam_System_RETURN (*)(PySpam_System_PROTO) PySpam_API[PySpam_System_NUM])

/* Return -1 on error, 0 on success.
 * PyCapsule_Import will set an exception if there's an error.
 */
static int
import_spam(void)
{
    PySpam_API = (void **)PyCapsule_Import("spam._C_API", 0);
    return (PySpam_API != NULL) ? 0 : -1;
}

#endif

#ifdef __cplusplus
}
#endif

#endif /* !defined(Py_SPAMMODULE_H) */

```

客户端模块必须在其初始化函数里按顺序调用函数 `import_spam()` (或其他宏) 才能访问函数 `PySpam_System()`。

```

PyMODINIT_FUNC
PyInit_client(void)
{
    PyObject *m;

    m = PyModule_Create(&clientmodule);
    if (m == NULL)
        return NULL;
    if (import_spam() < 0)
        return NULL;
    /* additional initialization can happen here */
}

```

(下页继续)

(续上页)

```

    return m;
}

```

这种方法的主要缺点是，文件 `spammodule.h` 过于复杂。当然，对每个要导出的函数，基本结构是相似的，所以只需要学习一次。

最后需要提醒的是 `Capsule` 提供了额外的功能，用于存储在 `Capsule` 里的指针的内存分配和释放。细节参考 `Python/C API 参考手册` 的章节 `capsules` 和 `Capsule` 的实现（在 `Python` 源码发行包的 `Include/pycapsule.h` 和 `Objects/pycapsule.c`）。

## 2.2 自定义扩展类型：教程

`Python` 允许编写 `C` 扩展模块定义可以从 `Python` 代码中操纵的新类型，这很像内置的 `str` 和 `list` 类型。所有扩展类型的代码都遵循一个模式，但是在您开始之前，您需要了解一些细节。这份文件是对这个主题介绍。

### 2.2.1 基础

`CPython` 运行时将所有 `Python` 对象都视为类型 `PyObject*` 的变量，即所有 `Python` 对象的“基础类型”。`PyObject` 结构体本身包含了对对象的 *reference count* 和对对象的“类型对象”。类型对象确定解释器需要调用哪些 (C) 函数，例如一个属性查询一个对象，一个方法调用，或者与另一个对象相乘。这些 `C` 函数被称为“类型方法”。

所以，如果你想要定义新的扩展类型，需要创建新的类型对象。

这类事情只能用例子解释，这里用一个最小化但完整的模块，定义了新的类型叫做 `Custom` 在 `C` 扩展模块 `custom` 里。

**注解：** 这里展示的方法是定义 *static* 扩展类型的传统方法。可以适合大部分用途。`C API` 也可以定义在堆上分配的扩展类型，使用 `PyType_FromSpec()` 函数，但不在本入门里讨论。

```

#define PY_SSIZE_T_CLEAN
#include <Python.h>

typedef struct {
    PyObject_HEAD
    /* Type-specific fields go here. */
} CustomObject;

static PyTypeObject CustomType = {
    PyVarObject_HEAD_INIT(NULL, 0)
    .tp_name = "custom.Custom",
    .tp_doc = "Custom objects",
    .tp_basicsize = sizeof(CustomObject),
    .tp_itemsize = 0,
    .tp_flags = Py_TPFLAGS_DEFAULT,
    .tp_new = PyType_GenericNew,
};

static PyModuleDef custommodule = {
    PyModuleDef_HEAD_INIT,
    .m_name = "custom",

```

(下页继续)

(续上页)

```

    .m_doc = "Example module that creates an extension type.",
    .m_size = -1,
};

PyMODINIT_FUNC
PyInit_custom(void)
{
    PyObject *m;
    if (PyType_Ready(&CustomType) < 0)
        return NULL;

    m = PyModule_Create(&custommodule);
    if (m == NULL)
        return NULL;

    Py_INCREF(&CustomType);
    PyModule_AddObject(m, "Custom", (PyObject *) &CustomType);
    return m;
}

```

这部分很容易理解，这是为了跟上一章能对接上。这个文件定义了三件事：

1. Custom 类的对象 **object** 包含了：CustomObject 结构，这会为每个 Custom 实例分配一次。
2. Custom **type** 的行为：这是 CustomType 结构体，其定义了一堆标识和函数指针，会指向解释器里请求的操作。
3. 初始化 custom 模块：PyInit\_custom 函数和对应的 custommodule 结构体。

结构的第一块是

```

typedef struct {
    PyObject_HEAD
} CustomObject;

```

这就是一个自定义对象所包含的。PyObject\_HEAD 是强制要求必须在每个对象结构体之前，用以定义一个类型为 PyObject 的字段叫 ob\_base，包含了一个指针指向类型对象和一个引用计数（这可以用宏 Py\_REFCNT 和 Py\_TYPE 来区分）。用宏来抽象，使得附加字段可以用调试构建。

**注解：**注意在宏 PyObject\_HEAD 后没有分号。意外添加分号会导致编译器提示出错。

当然，对象除了在 PyObject\_HEAD 存储数据外，还有额外数据；例如，如下定义了标准的 Python 浮点数：

```

typedef struct {
    PyObject_HEAD
    double ob_fval;
} PyFloatObject;

```

第二个位是类型对象的定义：

```

static PyTypeObject CustomType = {
    PyVarObject_HEAD_INIT(NULL, 0)
    .tp_name = "custom.Custom",
    .tp_doc = "Custom objects",
    .tp_basicsize = sizeof(CustomObject),
    .tp_itemsize = 0,

```

(下页继续)

(续上页)

```
.tp_flags = Py_TPFLAGS_DEFAULT,
.tp_new = PyType_GenericNew,
};
```

**注解：** 推荐使用如上 C99 风格的初始化，以避免列出所有的 `PyTypeObject` 字段，其中很多是你不需要关心的，这样也可以避免关注字段的定义顺序。

The actual definition of `PyTypeObject` in `object.h` has many more fields than the definition above. The remaining fields will be filled with zeros by the C compiler, and it's common practice to not specify them explicitly unless you need them.

We're going to pick it apart, one field at a time:

```
PyVarObject_HEAD_INIT(NULL, 0)
```

This line is mandatory boilerplate to initialize the `ob_base` field mentioned above.

```
.tp_name = "custom.Custom",
```

The name of our type. This will appear in the default textual representation of our objects and in some error messages, for example:

```
>>> "" + custom.Custom()
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
TypeError: can only concatenate str (not "custom.Custom") to str
```

Note that the name is a dotted name that includes both the module name and the name of the type within the module. The module in this case is `custom` and the type is `Custom`, so we set the type name to `custom.Custom`. Using the real dotted import path is important to make your type compatible with the `pydoc` and `pickle` modules.

```
.tp_basicsize = sizeof(CustomObject),
.tp_itemsize = 0,
```

This is so that Python knows how much memory to allocate when creating new `Custom` instances. `tp_itemsize` is only used for variable-sized objects and should otherwise be zero.

**注解：** If you want your type to be subclassable from Python, and your type has the same `tp_basicsize` as its base type, you may have problems with multiple inheritance. A Python subclass of your type will have to list your type first in its `__bases__`, or else it will not be able to call your type's `__new__()` method without getting an error. You can avoid this problem by ensuring that your type has a larger value for `tp_basicsize` than its base type does. Most of the time, this will be true anyway, because either your base type will be `object`, or else you will be adding data members to your base type, and therefore increasing its size.

We set the class flags to `Py_TPFLAGS_DEFAULT`.

```
.tp_flags = Py_TPFLAGS_DEFAULT,
```

All types should include this constant in their flags. It enables all of the members defined until at least Python 3.3. If you need further members, you will need to OR the corresponding flags.

We provide a doc string for the type in `tp_doc`.

```
.tp_doc = "Custom objects",
```

To enable object creation, we have to provide a `tp_new` handler. This is the equivalent of the Python method `__new__()`, but has to be specified explicitly. In this case, we can just use the default implementation provided by the API function `PyType_GenericNew()`.

```
.tp_new = PyType_GenericNew,
```

Everything else in the file should be familiar, except for some code in `PyInit_custom()`:

```
if (PyType_Ready(&CustomType) < 0)
    return;
```

This initializes the `Custom` type, filling in a number of members to the appropriate default values, including `ob_type` that we initially set to `NULL`.

```
PyModule_AddObject(m, "Custom", (PyObject *) &CustomType);
```

This adds the type to the module dictionary. This allows us to create `Custom` instances by calling the `Custom` class:

```
>>> import custom
>>> mycustom = custom.Custom()
```

That's it! All that remains is to build it; put the above code in a file called `custom.c` and:

```
from distutils.core import setup, Extension
setup(name="custom", version="1.0",
      ext_modules=[Extension("custom", ["custom.c"])])
```

in a file called `setup.py`; then typing

```
$ python setup.py build
```

at a shell should produce a file `custom.so` in a subdirectory; move to that directory and fire up Python — you should be able to import `custom` and play around with `Custom` objects.

That wasn't so hard, was it?

Of course, the current `Custom` type is pretty uninteresting. It has no data and doesn't do anything. It can't even be subclassed.

---

**注解:** While this documentation showcases the standard `distutils` module for building C extensions, it is recommended in real-world use cases to use the newer and better-maintained `setuptools` library. Documentation on how to do this is out of scope for this document and can be found in the [Python Packaging User's Guide](#).

---

## 2.2.2 Adding data and methods to the Basic example

Let's extend the basic example to add some data and methods. Let's also make the type usable as a base class. We'll create a new module, `custom2` that adds these capabilities:

```
#define PY_SSIZE_T_CLEAN
#include <Python.h>
#include "structmember.h"
```

(下页继续)

(续上页)

```

typedef struct {
    PyObject_HEAD
    PyObject *first; /* first name */
    PyObject *last;  /* last name */
    int number;
} CustomObject;

static void
Custom_dealloc(CustomObject *self)
{
    Py_XDECREF(self->first);
    Py_XDECREF(self->last);
    Py_TYPE(self)->tp_free((PyObject *) self);
}

static PyObject *
Custom_new(PyTypeObject *type, PyObject *args, PyObject *kwds)
{
    CustomObject *self;
    self = (CustomObject *) type->tp_alloc(type, 0);
    if (self != NULL) {
        self->first = PyUnicode_FromString("");
        if (self->first == NULL) {
            Py_DECREF(self);
            return NULL;
        }
        self->last = PyUnicode_FromString("");
        if (self->last == NULL) {
            Py_DECREF(self);
            return NULL;
        }
        self->number = 0;
    }
    return (PyObject *) self;
}

static int
Custom_init(CustomObject *self, PyObject *args, PyObject *kwds)
{
    static char *kwlist[] = {"first", "last", "number", NULL};
    PyObject *first = NULL, *last = NULL, *tmp;

    if (!PyArg_ParseTupleAndKeywords(args, kwds, "|OOi", kwlist,
                                     &first, &last,
                                     &self->number))
        return -1;

    if (first) {
        tmp = self->first;
        Py_INCREF(first);
        self->first = first;
        Py_XDECREF(tmp);
    }
    if (last) {
        tmp = self->last;
        Py_INCREF(last);
    }
}

```

(下页继续)

(续上页)

```

        self->last = last;
        Py_XDECREF(tmp);
    }
    return 0;
}

static PyMemberDef Custom_members[] = {
    {"first", T_OBJECT_EX, offsetof(CustomObject, first), 0,
     "first name"},
    {"last", T_OBJECT_EX, offsetof(CustomObject, last), 0,
     "last name"},
    {"number", T_INT, offsetof(CustomObject, number), 0,
     "custom number"},
    {NULL} /* Sentinel */
};

static PyObject *
Custom_name(CustomObject *self, PyObject *Py_UNUSED(ignored))
{
    if (self->first == NULL) {
        PyErr_SetString(PyExc_AttributeError, "first");
        return NULL;
    }
    if (self->last == NULL) {
        PyErr_SetString(PyExc_AttributeError, "last");
        return NULL;
    }
    return PyUnicode_FromFormat("%S %S", self->first, self->last);
}

static PyMethodDef Custom_methods[] = {
    {"name", (PyCFunction) Custom_name, METH_NOARGS,
     "Return the name, combining the first and last name"},
    {NULL} /* Sentinel */
};

static PyTypeObject CustomType = {
    PyVarObject_HEAD_INIT(NULL, 0)
    .tp_name = "custom2.Custom",
    .tp_doc = "Custom objects",
    .tp_basicsize = sizeof(CustomObject),
    .tp_itemsize = 0,
    .tp_flags = Py_TPFLAGS_DEFAULT | Py_TPFLAGS_BASETYPE,
    .tp_new = Custom_new,
    .tp_init = (initproc) Custom_init,
    .tp_dealloc = (destructor) Custom_dealloc,
    .tp_members = Custom_members,
    .tp_methods = Custom_methods,
};

static PyModuleDef custommodule = {
    PyModuleDef_HEAD_INIT,
    .m_name = "custom2",
    .m_doc = "Example module that creates an extension type.",
    .m_size = -1,

```

(下页继续)

(续上页)

```
};

PyMODINIT_FUNC
PyInit_custom2(void)
{
    PyObject *m;
    if (PyType_Ready(&CustomType) < 0)
        return NULL;

    m = PyModule_Create(&custommodule);
    if (m == NULL)
        return NULL;

    Py_INCREF(&CustomType);
    PyModule_AddObject(m, "Custom", (PyObject *) &CustomType);
    return m;
}
```

This version of the module has a number of changes.

We've added an extra include:

```
#include <structmember.h>
```

This include provides declarations that we use to handle attributes, as described a bit later.

The `Custom` type now has three data attributes in its C struct, *first*, *last*, and *number*. The *first* and *last* variables are Python strings containing first and last names. The *number* attribute is a C integer.

The object structure is updated accordingly:

```
typedef struct {
    PyObject_HEAD
    PyObject *first; /* first name */
    PyObject *last; /* last name */
    int number;
} CustomObject;
```

Because we now have data to manage, we have to be more careful about object allocation and deallocation. At a minimum, we need a deallocation method:

```
static void
Custom_dealloc(CustomObject *self)
{
    Py_XDECREF(self->first);
    Py_XDECREF(self->last);
    Py_TYPE(self)->tp_free((PyObject *) self);
}
```

which is assigned to the `tp_dealloc` member:

```
.tp_dealloc = (destructor) Custom_dealloc,
```

This method first clears the reference counts of the two Python attributes. `Py_XDECREF()` correctly handles the case where its argument is `NULL` (which might happen here if `tp_new` failed midway). It then calls the `tp_free` member of the object's type (computed by `Py_TYPE(self)`) to free the object's memory. Note that the object's type might not be `CustomType`, because the object may be an instance of a subclass.



**注解:** The explicit cast to destructor above is needed because we defined `Custom_dealloc` to take a `CustomObject *` argument, but the `tp_dealloc` function pointer expects to receive a `PyObject *` argument. Otherwise, the compiler will emit a warning. This is object-oriented polymorphism, in C!

We want to make sure that the first and last names are initialized to empty strings, so we provide a `tp_new` implementation:

```
static PyObject *
Custom_new(PyTypeObject *type, PyObject *args, PyObject *kwargs)
{
    CustomObject *self;
    self = (CustomObject *) type->tp_alloc(type, 0);
    if (self != NULL) {
        self->first = PyUnicode_FromString("");
        if (self->first == NULL) {
            Py_DECREF(self);
            return NULL;
        }
        self->last = PyUnicode_FromString("");
        if (self->last == NULL) {
            Py_DECREF(self);
            return NULL;
        }
        self->number = 0;
    }
    return (PyObject *) self;
}
```

and install it in the `tp_new` member:

```
.tp_new = Custom_new,
```

The `tp_new` handler is responsible for creating (as opposed to initializing) objects of the type. It is exposed in Python as the `__new__()` method. It is not required to define a `tp_new` member, and indeed many extension types will simply reuse `PyType_GenericNew()` as done in the first version of the `Custom` type above. In this case, we use the `tp_new` handler to initialize the `first` and `last` attributes to non-`NULL` default values.

`tp_new` is passed the type being instantiated (not necessarily `CustomType`, if a subclass is instantiated) and any arguments passed when the type was called, and is expected to return the instance created. `tp_new` handlers always accept positional and keyword arguments, but they often ignore the arguments, leaving the argument handling to initializer (a.k.a. `tp_init` in C or `__init__` in Python) methods.

**注解:** `tp_new` shouldn't call `tp_init` explicitly, as the interpreter will do it itself.

The `tp_new` implementation calls the `tp_alloc` slot to allocate memory:

```
self = (CustomObject *) type->tp_alloc(type, 0);
```

Since memory allocation may fail, we must check the `tp_alloc` result against `NULL` before proceeding.

**注解:** We didn't fill the `tp_alloc` slot ourselves. Rather `PyType_Ready()` fills it for us by inheriting it from our base class, which is `object` by default. Most types use the default allocation strategy.

**注解:** If you are creating a co-operative `tp_new` (one that calls a base type's `tp_new` or `__new__()`), you must *not* try to determine what method to call using method resolution order at runtime. Always statically determine what type you are going to call, and call its `tp_new` directly, or via `type->tp_base->tp_new`. If you do not do this, Python subclasses of your type that also inherit from other Python-defined classes may not work correctly. (Specifically, you may not be able to create instances of such subclasses without getting a `TypeError`.)

We also define an initialization function which accepts arguments to provide initial values for our instance:

```
static int
Custom_init(CustomObject *self, PyObject *args, PyObject *kwds)
{
    static char *kwlist[] = {"first", "last", "number", NULL};
    PyObject *first = NULL, *last = NULL, *tmp;

    if (!PyArg_ParseTupleAndKeywords(args, kwds, "|OOi", kwlist,
                                     &first, &last,
                                     &self->number))
        return -1;

    if (first) {
        tmp = self->first;
        Py_INCREF(first);
        self->first = first;
        Py_XDECREF(tmp);
    }
    if (last) {
        tmp = self->last;
        Py_INCREF(last);
        self->last = last;
        Py_XDECREF(tmp);
    }
    return 0;
}
```

by filling the `tp_init` slot.

```
.tp_init = (initproc) Custom_init,
```

The `tp_init` slot is exposed in Python as the `__init__()` method. It is used to initialize an object after it's created. Initializers always accept positional and keyword arguments, and they should return either 0 on success or -1 on error.

Unlike the `tp_new` handler, there is no guarantee that `tp_init` is called at all (for example, the `pickle` module by default doesn't call `__init__()` on unpickled instances). It can also be called multiple times. Anyone can call the `__init__()` method on our objects. For this reason, we have to be extra careful when assigning the new attribute values. We might be tempted, for example to assign the `first` member like this:

```
if (first) {
    Py_XDECREF(self->first);
    Py_INCREF(first);
    self->first = first;
}
```

But this would be risky. Our type doesn't restrict the type of the `first` member, so it could be any kind of object. It could have a destructor that causes code to be executed that tries to access the `first` member; or that destructor could release the *Global interpreter Lock* and let arbitrary code run in other threads that accesses and modifies our object.

To be paranoid and protect ourselves against this possibility, we almost always reassign members before decrementing their reference counts. When don't we have to do this?

- when we absolutely know that the reference count is greater than 1;
- when we know that deallocation of the object<sup>1</sup> will neither release the *GIL* nor cause any calls back into our type's code;
- when decrementing a reference count in a `tp_dealloc` handler on a type which doesn't support cyclic garbage collection<sup>2</sup>.

We want to expose our instance variables as attributes. There are a number of ways to do that. The simplest way is to define member definitions:

```
static PyMemberDef Custom_members[] = {
    {"first", T_OBJECT_EX, offsetof(CustomObject, first), 0,
     "first name"},
    {"last", T_OBJECT_EX, offsetof(CustomObject, last), 0,
     "last name"},
    {"number", T_INT, offsetof(CustomObject, number), 0,
     "custom number"},
    {NULL} /* Sentinel */
};
```

and put the definitions in the `tp_members` slot:

```
.tp_members = Custom_members,
```

Each member definition has a member name, type, offset, access flags and documentation string. See the [泛型属性管理](#) section below for details.

A disadvantage of this approach is that it doesn't provide a way to restrict the types of objects that can be assigned to the Python attributes. We expect the first and last names to be strings, but any Python objects can be assigned. Further, the attributes can be deleted, setting the C pointers to *NULL*. Even though we can make sure the members are initialized to non-*NULL* values, the members can be set to *NULL* if the attributes are deleted.

We define a single method, `Custom.name()`, that outputs the objects name as the concatenation of the first and last names.

```
static PyObject *
Custom_name(CustomObject *self, PyObject *Py_UNUSED(ignored))
{
    if (self->first == NULL) {
        PyErr_SetString(PyExc_AttributeError, "first");
        return NULL;
    }
    if (self->last == NULL) {
        PyErr_SetString(PyExc_AttributeError, "last");
        return NULL;
    }
    return PyUnicode_FromFormat("%S %S", self->first, self->last);
}
```

The method is implemented as a C function that takes a `Custom` (or `Custom` subclass) instance as the first argument. Methods always take an instance as the first argument. Methods often take positional and keyword arguments as well, but in this case we don't take any and don't need to accept a positional argument tuple or keyword argument dictionary. This method is equivalent to the Python method:

<sup>1</sup> This is true when we know that the object is a basic type, like a string or a float.

<sup>2</sup> We relied on this in the `tp_dealloc` handler in this example, because our type doesn't support garbage collection.

```
def name(self):
    return "%s %s" % (self.first, self.last)
```

Note that we have to check for the possibility that our `first` and `last` members are `NULL`. This is because they can be deleted, in which case they are set to `NULL`. It would be better to prevent deletion of these attributes and to restrict the attribute values to be strings. We'll see how to do that in the next section.

Now that we've defined the method, we need to create an array of method definitions:

```
static PyMethodDef Custom_methods[] = {
    {"name", (PyCFunction) Custom_name, METH_NOARGS,
     "Return the name, combining the first and last name"
    },
    {NULL} /* Sentinel */
};
```

(note that we used the `METH_NOARGS` flag to indicate that the method is expecting no arguments other than `self`)

and assign it to the `tp_methods` slot:

```
.tp_methods = Custom_methods,
```

Finally, we'll make our type usable as a base class for subclassing. We've written our methods carefully so far so that they don't make any assumptions about the type of the object being created or used, so all we need to do is to add the `Py_TPFLAGS_BASETYPE` to our class flag definition:

```
.tp_flags = Py_TPFLAGS_DEFAULT | Py_TPFLAGS_BASETYPE,
```

We rename `PyInit_custom()` to `PyInit_custom2()`, update the module name in the `PyModuleDef` struct, and update the full class name in the `PyTypeObject` struct.

Finally, we update our `setup.py` file to build the new module:

```
from distutils.core import setup, Extension
setup(name="custom", version="1.0",
      ext_modules=[
          Extension("custom", ["custom.c"]),
          Extension("custom2", ["custom2.c"]),
      ])
```

### 2.2.3 Providing finer control over data attributes

In this section, we'll provide finer control over how the `first` and `last` attributes are set in the `Custom` example. In the previous version of our module, the instance variables `first` and `last` could be set to non-string values or even deleted. We want to make sure that these attributes always contain strings.

```
#define PY_SSIZE_T_CLEAN
#include <Python.h>
#include "structmember.h"

typedef struct {
    PyObject_HEAD
    PyObject *first; /* first name */
    PyObject *last; /* last name */
    int number;
} CustomObject;
```

(下页继续)

(续上页)

```

static void
Custom_dealloc(CustomObject *self)
{
    Py_XDECREF(self->first);
    Py_XDECREF(self->last);
    Py_TYPE(self)->tp_free((PyObject *) self);
}

static PyObject *
Custom_new(PyTypeObject *type, PyObject *args, PyObject *kwargs)
{
    CustomObject *self;
    self = (CustomObject *) type->tp_alloc(type, 0);
    if (self != NULL) {
        self->first = PyUnicode_FromString("");
        if (self->first == NULL) {
            Py_DECREF(self);
            return NULL;
        }
        self->last = PyUnicode_FromString("");
        if (self->last == NULL) {
            Py_DECREF(self);
            return NULL;
        }
        self->number = 0;
    }
    return (PyObject *) self;
}

static int
Custom_init(CustomObject *self, PyObject *args, PyObject *kwargs)
{
    static char *kwlist[] = {"first", "last", "number", NULL};
    PyObject *first = NULL, *last = NULL, *tmp;

    if (!PyArg_ParseTupleAndKeywords(args, kwargs, "|UUii", kwlist,
                                     &first, &last,
                                     &self->number))
        return -1;

    if (first) {
        tmp = self->first;
        Py_INCREF(first);
        self->first = first;
        Py_DECREF(tmp);
    }
    if (last) {
        tmp = self->last;
        Py_INCREF(last);
        self->last = last;
        Py_DECREF(tmp);
    }
    return 0;
}

```

(下页继续)

(续上页)

```

static PyMemberDef Custom_members[] = {
    {"number", T_INT, offsetof(CustomObject, number), 0,
     "custom number"},
    {NULL} /* Sentinel */
};

static PyObject *
Custom_getfirst(CustomObject *self, void *closure)
{
    Py_INCREF(self->first);
    return self->first;
}

static int
Custom_setfirst(CustomObject *self, PyObject *value, void *closure)
{
    PyObject *tmp;
    if (value == NULL) {
        PyErr_SetString(PyExc_TypeError, "Cannot delete the first attribute");
        return -1;
    }
    if (!PyUnicode_Check(value)) {
        PyErr_SetString(PyExc_TypeError,
            "The first attribute value must be a string");
        return -1;
    }
    tmp = self->first;
    Py_INCREF(value);
    self->first = value;
    Py_DECREF(tmp);
    return 0;
}

static PyObject *
Custom_getlast(CustomObject *self, void *closure)
{
    Py_INCREF(self->last);
    return self->last;
}

static int
Custom_setlast(CustomObject *self, PyObject *value, void *closure)
{
    PyObject *tmp;
    if (value == NULL) {
        PyErr_SetString(PyExc_TypeError, "Cannot delete the last attribute");
        return -1;
    }
    if (!PyUnicode_Check(value)) {
        PyErr_SetString(PyExc_TypeError,
            "The last attribute value must be a string");
        return -1;
    }
    tmp = self->last;
    Py_INCREF(value);
    self->last = value;

```

(下页继续)

(续上页)

```

    Py_DECREF(tmp);
    return 0;
}

static PyGetSetDef Custom_getsetters[] = {
    {"first", (getter) Custom_getfirst, (setter) Custom_setfirst,
     "first name", NULL},
    {"last", (getter) Custom_getlast, (setter) Custom_setlast,
     "last name", NULL},
    {NULL} /* Sentinel */
};

static PyObject *
Custom_name(CustomObject *self, PyObject *Py_UNUSED(ignored))
{
    return PyUnicode_FromFormat("%S %S", self->first, self->last);
}

static PyMethodDef Custom_methods[] = {
    {"name", (PyCFunction) Custom_name, METH_NOARGS,
     "Return the name, combining the first and last name"},
    {NULL} /* Sentinel */
};

static PyTypeObject CustomType = {
    PyVarObject_HEAD_INIT(NULL, 0)
    .tp_name = "custom3.Custom",
    .tp_doc = "Custom objects",
    .tp_basicsize = sizeof(CustomObject),
    .tp_itemsize = 0,
    .tp_flags = Py_TPFLAGS_DEFAULT | Py_TPFLAGS_BASETYPE,
    .tp_new = Custom_new,
    .tp_init = (initproc) Custom_init,
    .tp_dealloc = (destructor) Custom_dealloc,
    .tp_members = Custom_members,
    .tp_methods = Custom_methods,
    .tp_getset = Custom_getsetters,
};

static PyModuleDef custommodule = {
    PyModuleDef_HEAD_INIT,
    .m_name = "custom3",
    .m_doc = "Example module that creates an extension type.",
    .m_size = -1,
};

PyMODINIT_FUNC
PyInit_custom3(void)
{
    PyObject *m;
    if (PyType_Ready(&CustomType) < 0)
        return NULL;

    m = PyModule_Create(&custommodule);
    if (m == NULL)

```

(下页继续)

(续上页)

```

        return NULL;

    Py_INCREF(&CustomType);
    PyModule_AddObject(m, "Custom", (PyObject *) &CustomType);
    return m;
}

```

To provide greater control, over the first and last attributes, we'll use custom getter and setter functions. Here are the functions for getting and setting the first attribute:

```

static PyObject *
Custom_getfirst(CustomObject *self, void *closure)
{
    Py_INCREF(self->first);
    return self->first;
}

static int
Custom_setfirst(CustomObject *self, PyObject *value, void *closure)
{
    PyObject *tmp;
    if (value == NULL) {
        PyErr_SetString(PyExc_TypeError, "Cannot delete the first attribute");
        return -1;
    }
    if (!PyUnicode_Check(value)) {
        PyErr_SetString(PyExc_TypeError,
                        "The first attribute value must be a string");
        return -1;
    }
    tmp = self->first;
    Py_INCREF(value);
    self->first = value;
    Py_DECREF(tmp);
    return 0;
}

```

The getter function is passed a `Custom` object and a "closure", which is a void pointer. In this case, the closure is ignored. (The closure supports an advanced usage in which definition data is passed to the getter and setter. This could, for example, be used to allow a single set of getter and setter functions that decide the attribute to get or set based on data in the closure.)

The setter function is passed the `Custom` object, the new value, and the closure. The new value may be `NULL`, in which case the attribute is being deleted. In our setter, we raise an error if the attribute is deleted or if its new value is not a string.

We create an array of `PyGetSetDef` structures:

```

static PyGetSetDef Custom_getsetters[] = {
    {"first", (getter) Custom_getfirst, (setter) Custom_setfirst,
     "first name", NULL},
    {"last", (getter) Custom_getlast, (setter) Custom_setlast,
     "last name", NULL},
    {NULL} /* Sentinel */
};

```

and register it in the `tp_getset` slot:



```
.tp_getset = Custom_getsetters,
```

The last item in a PyGetSetDef structure is the "closure" mentioned above. In this case, we aren't using a closure, so we just pass *NULL*.

We also remove the member definitions for these attributes:

```
static PyMemberDef Custom_members[] = {
    {"number", T_INT, offsetof(CustomObject, number), 0,
     "custom number"},
    {NULL} /* Sentinel */
};
```

We also need to update the `tp_init` handler to only allow strings<sup>3</sup> to be passed:

```
static int
Custom_init(CustomObject *self, PyObject *args, PyObject *kwargs)
{
    static char *kwlist[] = {"first", "last", "number", NULL};
    PyObject *first = NULL, *last = NULL, *tmp;

    if (!PyArg_ParseTupleAndKeywords(args, kwargs, "|UUi", kwlist,
                                     &first, &last,
                                     &self->number))
        return -1;

    if (first) {
        tmp = self->first;
        Py_INCREF(first);
        self->first = first;
        Py_DECREF(tmp);
    }
    if (last) {
        tmp = self->last;
        Py_INCREF(last);
        self->last = last;
        Py_DECREF(tmp);
    }
    return 0;
}
```

With these changes, we can assure that the `first` and `last` members are never *NULL* so we can remove checks for *NULL* values in almost all cases. This means that most of the `Py_XDECREF()` calls can be converted to `Py_DECREF()` calls. The only place we can't change these calls is in the `tp_dealloc` implementation, where there is the possibility that the initialization of these members failed in `tp_new`.

We also rename the module initialization function and module name in the initialization function, as we did before, and we add an extra definition to the `setup.py` file.

## 2.2.4 Supporting cyclic garbage collection

Python has a *cyclic garbage collector (GC)* that can identify unneeded objects even when their reference counts are not zero. This can happen when objects are involved in cycles. For example, consider:

<sup>3</sup> We now know that the `first` and `last` members are strings, so perhaps we could be less careful about decrementing their reference counts, however, we accept instances of string subclasses. Even though deallocating normal strings won't call back into our objects, we can't guarantee that deallocating an instance of a string subclass won't call back into our objects.

```
>>> l = []
>>> l.append(l)
>>> del l
```

In this example, we create a list that contains itself. When we delete it, it still has a reference from itself. Its reference count doesn't drop to zero. Fortunately, Python's cyclic garbage collector will eventually figure out that the list is garbage and free it.

In the second version of the `Custom` example, we allowed any kind of object to be stored in the `first` or `last` attributes<sup>4</sup>. Besides, in the second and third versions, we allowed subclassing `Custom`, and subclasses may add arbitrary attributes. For any of those two reasons, `Custom` objects can participate in cycles:

```
>>> import custom3
>>> class Derived(custom3.Custom): pass
...
>>> n = Derived()
>>> n.some_attribute = n
```

To allow a `Custom` instance participating in a reference cycle to be properly detected and collected by the cyclic GC, our `Custom` type needs to fill two additional slots and to enable a flag that enables these slots:

```
#define PY_SSIZE_T_CLEAN
#include <Python.h>
#include "structmember.h"

typedef struct {
    PyObject_HEAD
    PyObject *first; /* first name */
    PyObject *last; /* last name */
    int number;
} CustomObject;

static int
Custom_traverse(CustomObject *self, visitproc visit, void *arg)
{
    Py_VISIT(self->first);
    Py_VISIT(self->last);
    return 0;
}

static int
Custom_clear(CustomObject *self)
{
    Py_CLEAR(self->first);
    Py_CLEAR(self->last);
    return 0;
}

static void
Custom_dealloc(CustomObject *self)
{
    PyObject_GC_UnTrack(self);
    Custom_clear(self);
    Py_TYPE(self)->tp_free((PyObject *) self);
}
```

(下页继续)

<sup>4</sup> Also, even with our attributes restricted to strings instances, the user could pass arbitrary `str` subclasses and therefore still create reference cycles.

(续上页)

```

static PyObject *
Custom_new(PyTypeObject *type, PyObject *args, PyObject *kwargs)
{
    CustomObject *self;
    self = (CustomObject *) type->tp_alloc(type, 0);
    if (self != NULL) {
        self->first = PyUnicode_FromString("");
        if (self->first == NULL) {
            Py_DECREF(self);
            return NULL;
        }
        self->last = PyUnicode_FromString("");
        if (self->last == NULL) {
            Py_DECREF(self);
            return NULL;
        }
        self->number = 0;
    }
    return (PyObject *) self;
}

static int
Custom_init(CustomObject *self, PyObject *args, PyObject *kwargs)
{
    static char *kwlist[] = {"first", "last", "number", NULL};
    PyObject *first = NULL, *last = NULL, *tmp;

    if (!PyArg_ParseTupleAndKeywords(args, kwargs, "|UUi", kwlist,
                                     &first, &last,
                                     &self->number))
        return -1;

    if (first) {
        tmp = self->first;
        Py_INCREF(first);
        self->first = first;
        Py_DECREF(tmp);
    }
    if (last) {
        tmp = self->last;
        Py_INCREF(last);
        self->last = last;
        Py_DECREF(tmp);
    }
    return 0;
}

static PyMemberDef Custom_members[] = {
    {"number", T_INT, offsetof(CustomObject, number), 0,
     "custom number"},
    {NULL} /* Sentinel */
};

static PyObject *
Custom_getfirst(CustomObject *self, void *closure)
{

```

(下页继续)

(续上页)

```

    Py_INCREF(self->first);
    return self->first;
}

static int
Custom_setfirst(CustomObject *self, PyObject *value, void *closure)
{
    if (value == NULL) {
        PyErr_SetString(PyExc_TypeError, "Cannot delete the first attribute");
        return -1;
    }
    if (!PyUnicode_Check(value)) {
        PyErr_SetString(PyExc_TypeError,
                        "The first attribute value must be a string");
        return -1;
    }
    Py_INCREF(value);
    Py_CLEAR(self->first);
    self->first = value;
    return 0;
}

static PyObject *
Custom_getlast(CustomObject *self, void *closure)
{
    Py_INCREF(self->last);
    return self->last;
}

static int
Custom_setlast(CustomObject *self, PyObject *value, void *closure)
{
    if (value == NULL) {
        PyErr_SetString(PyExc_TypeError, "Cannot delete the last attribute");
        return -1;
    }
    if (!PyUnicode_Check(value)) {
        PyErr_SetString(PyExc_TypeError,
                        "The last attribute value must be a string");
        return -1;
    }
    Py_INCREF(value);
    Py_CLEAR(self->last);
    self->last = value;
    return 0;
}

static PyGetSetDef Custom_getsetters[] = {
    {"first", (getter) Custom_getfirst, (setter) Custom_setfirst,
     "first name", NULL},
    {"last", (getter) Custom_getlast, (setter) Custom_setlast,
     "last name", NULL},
    {NULL} /* Sentinel */
};

static PyObject *

```

(下页继续)

(续上页)

```

Custom_name(CustomObject *self, PyObject *Py_UNUSED(ignored))
{
    return PyUnicode_FromFormat("%S %S", self->first, self->last);
}

static PyMethodDef Custom_methods[] = {
    {"name", (PyCFunction) Custom_name, METH_NOARGS,
     "Return the name, combining the first and last name"
    },
    {NULL} /* Sentinel */
};

static PyTypeObject CustomType = {
    PyVarObject_HEAD_INIT(NULL, 0)
    .tp_name = "custom4.Custom",
    .tp_doc = "Custom objects",
    .tp_basicsize = sizeof(CustomObject),
    .tp_itemsize = 0,
    .tp_flags = Py_TPFLAGS_DEFAULT | Py_TPFLAGS_BASETYPE | Py_TPFLAGS_HAVE_GC,
    .tp_new = Custom_new,
    .tp_init = (initproc) Custom_init,
    .tp_dealloc = (destructor) Custom_dealloc,
    .tp_traverse = (traverseproc) Custom_traverse,
    .tp_clear = (inquiry) Custom_clear,
    .tp_members = Custom_members,
    .tp_methods = Custom_methods,
    .tp_getset = Custom_getsetters,
};

static PyModuleDef custommodule = {
    PyModuleDef_HEAD_INIT,
    .m_name = "custom4",
    .m_doc = "Example module that creates an extension type.",
    .m_size = -1,
};

PyMODINIT_FUNC
PyInit_custom4(void)
{
    PyObject *m;
    if (PyType_Ready(&CustomType) < 0)
        return NULL;

    m = PyModule_Create(&custommodule);
    if (m == NULL)
        return NULL;

    Py_INCREF(&CustomType);
    PyModule_AddObject(m, "Custom", (PyObject *) &CustomType);
    return m;
}

```

First, the traversal method lets the cyclic GC know about subobjects that could participate in cycles:

```

static int
Custom_traverse(CustomObject *self, visitproc visit, void *arg)

```

(下页继续)

(续上页)

```
{
    int vret;
    if (self->first) {
        vret = visit(self->first, arg);
        if (vret != 0)
            return vret;
    }
    if (self->last) {
        vret = visit(self->last, arg);
        if (vret != 0)
            return vret;
    }
    return 0;
}
```

For each subobject that can participate in cycles, we need to call the `visit()` function, which is passed to the traversal method. The `visit()` function takes as arguments the subobject and the extra argument `arg` passed to the traversal method. It returns an integer value that must be returned if it is non-zero.

Python provides a `Py_VISIT()` macro that automates calling visit functions. With `Py_VISIT()`, we can minimize the amount of boilerplate in `Custom_traverse`:

```
static int
Custom_traverse(CustomObject *self, visitproc visit, void *arg)
{
    Py_VISIT(self->first);
    Py_VISIT(self->last);
    return 0;
}
```

**注解:** The `tp_traverse` implementation must name its arguments exactly `visit` and `arg` in order to use `Py_VISIT()`.

Second, we need to provide a method for clearing any subobjects that can participate in cycles:

```
static int
Custom_clear(CustomObject *self)
{
    Py_CLEAR(self->first);
    Py_CLEAR(self->last);
    return 0;
}
```

Notice the use of the `Py_CLEAR()` macro. It is the recommended and safe way to clear data attributes of arbitrary types while decrementing their reference counts. If you were to call `Py_XDECREF()` instead on the attribute before setting it to `NULL`, there is a possibility that the attribute's destructor would call back into code that reads the attribute again (*especially* if there is a reference cycle).

**注解:** You could emulate `Py_CLEAR()` by writing:

```
PyObject *tmp;
tmp = self->first;
self->first = NULL;
Py_XDECREF(tmp);
```

Nevertheless, it is much easier and less error-prone to always use `Py_CLEAR()` when deleting an attribute. Don't try to micro-optimize at the expense of robustness!

The deallocator `Custom_dealloc` may call arbitrary code when clearing attributes. It means the circular GC can be triggered inside the function. Since the GC assumes reference count is not zero, we need to untrack the object from the GC by calling `PyObject_GC_UnTrack()` before clearing members. Here is our reimplemented deallocator using `PyObject_GC_UnTrack()` and `Custom_clear`:

```
static void
Custom_dealloc(CustomObject *self)
{
    PyObject_GC_UnTrack(self);
    Custom_clear(self);
    Py_TYPE(self)->tp_free((PyObject *) self);
}
```

Finally, we add the `Py_TPFLAGS_HAVE_GC` flag to the class flags:

```
.tp_flags = Py_TPFLAGS_DEFAULT | Py_TPFLAGS_BASETYPE | Py_TPFLAGS_HAVE_GC,
```

That's pretty much it. If we had written custom `tp_alloc` or `tp_free` handlers, we'd need to modify them for cyclic garbage collection. Most extensions will use the versions automatically provided.

## 2.2.5 Subclassing other types

It is possible to create new extension types that are derived from existing types. It is easiest to inherit from the built in types, since an extension can easily use the `PyTypeObject` it needs. It can be difficult to share these `PyTypeObject` structures between extension modules.

In this example we will create a `SubList` type that inherits from the built-in `list` type. The new type will be completely compatible with regular lists, but will have an additional `increment()` method that increases an internal counter:

```
>>> import sublist
>>> s = sublist.SubList(range(3))
>>> s.extend(s)
>>> print(len(s))
6
>>> print(s.increment())
1
>>> print(s.increment())
2
```

```
#define PY_SSIZE_T_CLEAN
#include <Python.h>

typedef struct {
    PyListObject list;
    int state;
} SubListObject;

static PyObject *
SubList_increment(SubListObject *self, PyObject *unused)
{
    self->state++;
    return PyLong_FromLong(self->state);
}
```

(下页继续)

(续上页)

```

}

static PyMethodDef SubList_methods[] = {
    {"increment", (PyCFunction) SubList_increment, METH_NOARGS,
     PyDoc_STR("increment state counter")},
    {NULL},
};

static int
SubList_init(SubListObject *self, PyObject *args, PyObject *kwds)
{
    if (PyList_Type.tp_init((PyObject *) self, args, kwds) < 0)
        return -1;
    self->state = 0;
    return 0;
}

static PyTypeObject SubListType = {
    PyVarObject_HEAD_INIT(NULL, 0)
    .tp_name = "sublist.SubList",
    .tp_doc = "SubList objects",
    .tp_basicsize = sizeof(SubListObject),
    .tp_itemsize = 0,
    .tp_flags = Py_TPFLAGS_DEFAULT | Py_TPFLAGS_BASETYPE,
    .tp_init = (initproc) SubList_init,
    .tp_methods = SubList_methods,
};

static PyModuleDef sublistmodule = {
    PyModuleDef_HEAD_INIT,
    .m_name = "sublist",
    .m_doc = "Example module that creates an extension type.",
    .m_size = -1,
};

PyMODINIT_FUNC
PyInit_sublist(void)
{
    PyObject *m;
    SubListType.tp_base = &PyList_Type;
    if (PyType_Ready(&SubListType) < 0)
        return NULL;

    m = PyModule_Create(&sublistmodule);
    if (m == NULL)
        return NULL;

    Py_INCREF(&SubListType);
    PyModule_AddObject(m, "SubList", (PyObject *) &SubListType);
    return m;
}
    
```

As you can see, the source code closely resembles the Custom examples in previous sections. We will break down the main differences between them.

```
typedef struct {
```

(下页继续)



(续上页)

```
PyListObject list;
    int state;
} SubListObject;
```

The primary difference for derived type objects is that the base type's object structure must be the first value. The base type will already include the `PyObject_HEAD()` at the beginning of its structure.

When a Python object is a `SubList` instance, its `PyObject *` pointer can be safely cast to both `PyListObject *` and `SubListObject *`:

```
static int
SubList_init(SubListObject *self, PyObject *args, PyObject *kwds)
{
    if (PyList_Type.tp_init((PyObject *) self, args, kwds) < 0)
        return -1;
    self->state = 0;
    return 0;
}
```

We see above how to call through to the `__init__` method of the base type.

This pattern is important when writing a type with custom `tp_new` and `tp_dealloc` members. The `tp_new` handler should not actually create the memory for the object with its `tp_alloc`, but let the base class handle it by calling its own `tp_new`.

The `PyTypeObject` struct supports a `tp_base` specifying the type's concrete base class. Due to cross-platform compiler issues, you can't fill that field directly with a reference to `PyList_Type`; it should be done later in the module initialization function:

```
PyMODINIT_FUNC
PyInit_sublist(void)
{
    PyObject* m;
    SubListType.tp_base = &PyList_Type;
    if (PyType_Ready(&SubListType) < 0)
        return NULL;

    m = PyModule_Create(&sublistmodule);
    if (m == NULL)
        return NULL;

    Py_INCREF(&SubListType);
    PyModule_AddObject(m, "SubList", (PyObject *) &SubListType);
    return m;
}
```

Before calling `PyType_Ready()`, the type structure must have the `tp_base` slot filled in. When we are deriving an existing type, it is not necessary to fill out the `tp_alloc` slot with `PyType_GenericNew()` – the allocation function from the base type will be inherited.

After that, calling `PyType_Ready()` and adding the type object to the module is the same as with the basic `Custom` examples.

## 2.3 定义扩展类型：已分类主题

本章节目标是提供一个各种你可以实现的类型方法及其功能的简短介绍。

这是 C 类型 PyTypeObject 的定义，省略了只用于调试构建的字段：

```
typedef struct _typeobject {
    PyObject_VAR_HEAD
    const char *tp_name; /* For printing, in format "<module>.<name>" */
    Py_ssize_t tp_basicsize, tp_itemsize; /* For allocation */

    /* Methods to implement standard operations */

    destructor tp_dealloc;
    Py_ssize_t tp_vectorcall_offset;
    getattrofunc tp_getattr;
    setattrofunc tp_setattr;
    PyAsyncMethods *tp_as_async; /* formerly known as tp_compare (Python 2)
                                   or tp_reserved (Python 3) */

    reprfunc tp_repr;

    /* Method suites for standard classes */

    PyNumberMethods *tp_as_number;
    PySequenceMethods *tp_as_sequence;
    PyMappingMethods *tp_as_mapping;

    /* More standard operations (here for binary compatibility) */

    hashfunc tp_hash;
    ternaryfunc tp_call;
    reprfunc tp_str;
    getattrofunc tp_getattro;
    setattrofunc tp_setattro;

    /* Functions to access object as input/output buffer */
    PyBufferProcs *tp_as_buffer;

    /* Flags to define presence of optional/expanded features */
    unsigned long tp_flags;

    const char *tp_doc; /* Documentation string */

    /* call function for all accessible objects */
    traverseproc tp_traverse;

    /* delete references to contained objects */
    inquiry tp_clear;

    /* rich comparisons */
    richcmpfunc tp_richcompare;

    /* weak reference enabler */
    Py_ssize_t tp_weaklistoffset;

    /* Iterators */
    getiterfunc tp_iter;
    iternextfunc tp_iternext;

    /* Attribute descriptor and subclassing stuff */
    struct PyMethodDef *tp_methods;
    struct PyMemberDef *tp_members;
```

(下页继续)

(续上页)

```

struct PyGetSetDef *tp_getset;
struct _typeobject *tp_base;
PyObject *tp_dict;
descrgetfunc tp_descr_get;
descrsetfunc tp_descr_set;
Py_ssize_t tp_dictoffset;
initproc tp_init;
allocfunc tp_alloc;
newfunc tp_new;
freefunc tp_free; /* Low-level free-memory routine */
inquiry tp_is_gc; /* For PyObject_IS_GC */
PyObject *tp_bases;
PyObject *tp_mro; /* method resolution order */
PyObject *tp_cache;
PyObject *tp_subclasses;
PyObject *tp_weaklist;
destructor tp_del;

/* Type attribute cache version tag. Added in version 2.6 */
unsigned int tp_version_tag;

destructor tp_finalize;
} PyTypeObject;

```

这里有很多方法。但是不要太担心，如果你要定义一个类型，通常只需要实现少量的方法。

正如你猜到的一样，我们正要一步一步详细介绍各种处理程序。因为有大量的历史包袱影响字段的排序，所以我们不会根据它们在结构体里定义的顺序讲解。通常非常容易找到一个包含你需要的字段的例子，然后改变值去适应你新的类型。

```
const char *tp_name; /* For printing */
```

类型的名字 - 上一章提到过的，会出现在很多地方，几乎全部都是为了诊断目的。尝试选择一个好名字，对于诊断很有帮助。

```
Py_ssize_t tp_basicsize, tp_itemsize; /* For allocation */
```

这些字段告诉运行时在创造这个类型的新对象时需要分配多少内存。Python 为了可变长度的结构（想下：字符串，元组）有些内置支持，这是 `tp_itemsize` 字段存在的原由。这部分稍后解释。

```
const char *tp_doc;
```

这里你可以放置一段字符串（或者它的地址），当你想在 Python 脚本引用 `obj.__doc__` 时返回这段文档字符串。

现在我们来看一下基本类型方法 - 大多数扩展类型将实现的方法。

### 2.3.1 终结和内存释放

```
destructor tp_dealloc;
```

当您的类型实例的引用计数减少为零并且 Python 解释器想要回收它时，将调用此函数。如果你的类型有内存可供释放或执行其他清理，你可以把它放在这里。对象本身也需要在这里释放。以下是此函数的示例：

```
static void
newdatatype_dealloc(newdatatypeobject *obj)
{
    free(obj->obj_UnderlyingDatatypePtr);
    Py_TYPE(obj)->tp_free(obj);
}
```

One important requirement of the deallocator function is that it leaves any pending exceptions alone. This is important since deallocators are frequently called as the interpreter unwinds the Python stack; when the stack is unwound due to an exception (rather than normal returns), nothing is done to protect the deallocators from seeing that an exception has already been set. Any actions which a deallocator performs which may cause additional Python code to be executed may detect that an exception has been set. This can lead to misleading errors from the interpreter. The proper way to protect against this is to save a pending exception before performing the unsafe action, and restoring it when done. This can be done using the `PyErr_Fetch()` and `PyErr_Restore()` functions:

```
static void
my_dealloc(PyObject *obj)
{
    PyObject *self = (PyObject *) obj;
    PyObject *cbresult;

    if (self->my_callback != NULL) {
        PyObject *err_type, *err_value, *err_traceback;

        /* This saves the current exception state */
        PyErr_Fetch(&err_type, &err_value, &err_traceback);

        cbresult = PyObject_CallObject(self->my_callback, NULL);
        if (cbresult == NULL)
            PyErr_WriteUnraisable(self->my_callback);
        else
            Py_DECREF(cbresult);

        /* This restores the saved exception state */
        PyErr_Restore(err_type, err_value, err_traceback);

        Py_DECREF(self->my_callback);
    }
    Py_TYPE(obj)->tp_free((PyObject*) self);
}
```

**注解:** There are limitations to what you can safely do in a deallocator function. First, if your type supports garbage collection (using `tp_traverse` and/or `tp_clear`), some of the object's members can have been cleared or finalized by the time `tp_dealloc` is called. Second, in `tp_dealloc`, your object is in an unstable state: its reference count is equal to zero. Any call to a non-trivial object or API (as in the example above) might end up calling `tp_dealloc` again, causing a double free and a crash.

从 Python 3.4 开始, 推荐不要在 `tp_dealloc` 放复杂的终结代码, 而是使用新的 `tp_finalize` 类型方法。

**参见:**

**PEP 442** 解释了新的终结方案。

### 2.3.2 对象展示

在 Python 中，有两种方式可以生成对象的文本表示: `repr()` 函数和 `str()` 函数。( `print()` 函数会直接调用 `str()`。 ) 这些处理程序都是可选的。

```
reprfunc tp_repr;
reprfunc tp_str;
```

`tp_repr` 处理程序应该返回一个字符串对象，其中包含调用它的实例的表示形式。下面是一个简单的例子:

```
static PyObject *
newdatatype_repr(newdatatypeobject * obj)
{
    return PyUnicode_FromFormat("Repr-ified_newdatatype{{size:%d}}",
                                obj->obj_UnderlyingDatatypePtr->size);
}
```

如果没有指定 `tp_repr` 处理程序，解释器将提供一个使用 `tp_name` 的表示形式以及对象的惟一标识值。

The `tp_str` handler is to `str()` what the `tp_repr` handler described above is to `repr()`; that is, it is called when Python code calls `str()` on an instance of your object. Its implementation is very similar to the `tp_repr` function, but the resulting string is intended for human consumption. If `tp_str` is not specified, the `tp_repr` handler is used instead.

下面是一个简单的例子:

```
static PyObject *
newdatatype_str(newdatatypeobject * obj)
{
    return PyUnicode_FromFormat("Stringified_newdatatype{{size:%d}}",
                                obj->obj_UnderlyingDatatypePtr->size);
}
```

### 2.3.3 Attribute Management

For every object which can support attributes, the corresponding type must provide the functions that control how the attributes are resolved. There needs to be a function which can retrieve attributes (if any are defined), and another to set attributes (if setting attributes is allowed). Removing an attribute is a special case, for which the new value passed to the handler is *NULL*.

Python supports two pairs of attribute handlers; a type that supports attributes only needs to implement the functions for one pair. The difference is that one pair takes the name of the attribute as a `char*`, while the other accepts a `PyObject*`. Each type can use whichever pair makes more sense for the implementation's convenience.

```
getattrfunc tp_getattr;          /* char * version */
setattrfunc tp_setattr;
/* ... */
getattrofunc tp_getattro;        /* PyObject * version */
setattrofunc tp_setattro;
```

If accessing attributes of an object is always a simple operation (this will be explained shortly), there are generic implementations which can be used to provide the `PyObject*` version of the attribute management functions. The actual need for type-specific attribute handlers almost completely disappeared starting with Python 2.2, though there are many examples which have not been updated to use some of the new generic mechanism that is available.

## 泛型属性管理

Most extension types only use *simple* attributes. So, what makes the attributes simple? There are only a couple of conditions that must be met:

1. The name of the attributes must be known when `PyType_Ready()` is called.
2. 不需要特殊的处理来记录属性是否被查找或设置，也不需要根据值采取操作。

请注意，此列表不对属性的值、值的计算时间或相关数据的存储方式施加任何限制。

When `PyType_Ready()` is called, it uses three tables referenced by the type object to create *descriptors* which are placed in the dictionary of the type object. Each descriptor controls access to one attribute of the instance object. Each of the tables is optional; if all three are `NULL`, instances of the type will only have attributes that are inherited from their base type, and should leave the `tp_getattro` and `tp_setattro` fields `NULL` as well, allowing the base type to handle attributes.

表被声明为 `object::` 类型的三个字段:

```
struct PyMethodDef *tp_methods;
struct PyMemberDef *tp_members;
struct PyGetSetDef *tp_getset;
```

如果 `tp_methods` 非 `NULL`，它必须引用一个 `PyMethodDef` 结构的数组。表中的每个条目都是这个结构的一个实例:

```
typedef struct PyMethodDef {
    const char *ml_name;           /* method name */
    PyCFunction ml_meth;           /* implementation function */
    int ml_flags;                  /* flags */
    const char *ml_doc;            /* docstring */
} PyMethodDef;
```

One entry should be defined for each method provided by the type; no entries are needed for methods inherited from a base type. One additional entry is needed at the end; it is a sentinel that marks the end of the array. The `ml_name` field of the sentinel must be `NULL`.

The second table is used to define attributes which map directly to data stored in the instance. A variety of primitive C types are supported, and access may be read-only or read-write. The structures in the table are defined as:

```
typedef struct PyMemberDef {
    const char *name;
    int type;
    int offset;
    int flags;
    const char *doc;
} PyMemberDef;
```

For each entry in the table, a *descriptor* will be constructed and added to the type which will be able to extract a value from the instance structure. The `type` field should contain one of the type codes defined in the `structmember.h` header; the value will be used to determine how to convert Python values to and from C values. The `flags` field is used to store flags which control how the attribute can be accessed.

以下标志常量定义在 `file: 'structmember.h'`; 它们可以使用 bitwise-OR 组合。

常数	意义
READONLY	没有可写的
READ_RESTRICTED	Not readable in restricted mode.
WRITE_RESTRICTED	Not writable in restricted mode.
RESTRICTED	在受限模式下不可读，也不可写。

An interesting advantage of using the `tp_members` table to build descriptors that are used at runtime is that any attribute defined this way can have an associated doc string simply by providing the text in the table. An application can use the introspection API to retrieve the descriptor from the class object, and get the doc string using its `__doc__` attribute.

As with the `tp_methods` table, a sentinel entry with a name value of `NULL` is required.

## Type-specific Attribute Management

For simplicity, only the `char*` version will be demonstrated here; the type of the name parameter is the only difference between the `char*` and `PyObject*` flavors of the interface. This example effectively does the same thing as the generic example above, but does not use the generic support added in Python 2.2. It explains how the handler functions are called, so that if you do need to extend their functionality, you'll understand what needs to be done.

The `tp_getattr` handler is called when the object requires an attribute look-up. It is called in the same situations where the `__getattr__()` method of a class would be called.

Here is an example:

```
static PyObject *
newdatatype_getattr(newdatatypeobject *obj, char *name)
{
    if (strcmp(name, "data") == 0)
    {
        return PyLong_FromLong(obj->data);
    }

    PyErr_Format(PyExc_AttributeError,
        "'%.50s' object has no attribute '%.400s'",
        tp->tp_name, name);
    return NULL;
}
```

The `tp_setattr` handler is called when the `__setattr__()` or `__delattr__()` method of a class instance would be called. When an attribute should be deleted, the third parameter will be `NULL`. Here is an example that simply raises an exception; if this were really all you wanted, the `tp_setattr` handler should be set to `NULL`.

```
static int
newdatatype_setattr(newdatatypeobject *obj, char *name, PyObject *v)
{
    PyErr_Format(PyExc_RuntimeError, "Read-only attribute: %s", name);
    return -1;
}
```

## 2.3.4 Object Comparison

```
richcmpfunc tp_richcompare;
```

The `tp_richcompare` handler is called when comparisons are needed. It is analogous to the rich comparison methods, like `__lt__()`, and also called by `PyObject_RichCompare()` and `PyObject_RichCompareBool()`.

This function is called with two Python objects and the operator as arguments, where the operator is one of `Py_EQ`, `Py_NE`, `Py_LE`, `Py_GT`, `Py_LT` or `Py_GE`. It should compare the two objects with respect to the specified operator and return `Py_True` or `Py_False` if the comparison is successful, `Py_NotImplemented` to indicate that comparison is not implemented and the other object's comparison method should be tried, or `NULL` if an exception was set.

Here is a sample implementation, for a datatype that is considered equal if the size of an internal pointer is equal:

```
static PyObject *
newdatatype_richcmp(PyObject *obj1, PyObject *obj2, int op)
{
    PyObject *result;
    int c, size1, size2;

    /* code to make sure that both arguments are of type
       newdatatype omitted */

    size1 = obj1->obj_UnderlyingDatatypePtr->size;
    size2 = obj2->obj_UnderlyingDatatypePtr->size;

    switch (op) {
        case Py_LT: c = size1 < size2; break;
        case Py_LE: c = size1 <= size2; break;
        case Py_EQ: c = size1 == size2; break;
        case Py_NE: c = size1 != size2; break;
        case Py_GT: c = size1 > size2; break;
        case Py_GE: c = size1 >= size2; break;
    }
    result = c ? Py_True : Py_False;
    Py_INCREF(result);
    return result;
}
```

## 2.3.5 Abstract Protocol Support

Python supports a variety of *abstract* 'protocols,' the specific interfaces provided to use these interfaces are documented in abstract.

A number of these abstract interfaces were defined early in the development of the Python implementation. In particular, the number, mapping, and sequence protocols have been part of Python since the beginning. Other protocols have been added over time. For protocols which depend on several handler routines from the type implementation, the older protocols have been defined as optional blocks of handlers referenced by the type object. For newer protocols there are additional slots in the main type object, with a flag bit being set to indicate that the slots are present and should be checked by the interpreter. (The flag bit does not indicate that the slot values are non-*NULL*. The flag may be set to indicate the presence of a slot, but a slot may still be unfilled.)

```
PyNumberMethods    *tp_as_number;
PySequenceMethods  *tp_as_sequence;
PyMappingMethods    *tp_as_mapping;
```

If you wish your object to be able to act like a number, a sequence, or a mapping object, then you place the address of a structure that implements the C type `PyNumberMethods`, `PySequenceMethods`, or `PyMappingMethods`, respectively. It is up to you to fill in this structure with appropriate values. You can find examples of the use of each of these in the `Objects` directory of the Python source distribution.



```
hashfunc tp_hash;
```

This function, if you choose to provide it, should return a hash number for an instance of your data type. Here is a simple example:

```
static Py_hash_t
newdatatype_hash(newdatatypeobject *obj)
{
    Py_hash_t result;
    result = obj->some_size + 32767 * obj->some_number;
    if (result == -1)
        result = -2;
    return result;
}
```

`Py_hash_t` is a signed integer type with a platform-varying width. Returning `-1` from `tp_hash` indicates an error, which is why you should be careful to avoid returning it when hash computation is successful, as seen above.

```
ternaryfunc tp_call;
```

This function is called when an instance of your data type is "called", for example, if `obj1` is an instance of your data type and the Python script contains `obj1('hello')`, the `tp_call` handler is invoked.

This function takes three arguments:

1. *self* is the instance of the data type which is the subject of the call. If the call is `obj1('hello')`, then *self* is `obj1`.
2. *args* is a tuple containing the arguments to the call. You can use `PyArg_ParseTuple()` to extract the arguments.
3. *kwds* is a dictionary of keyword arguments that were passed. If this is non-*NULL* and you support keyword arguments, use `PyArg_ParseTupleAndKeywords()` to extract the arguments. If you do not want to support keyword arguments and this is non-*NULL*, raise a `TypeError` with a message saying that keyword arguments are not supported.

Here is a toy `tp_call` implementation:

```
static PyObject *
newdatatype_call(newdatatypeobject *self, PyObject *args, PyObject *kwds)
{
    PyObject *result;
    const char *arg1;
    const char *arg2;
    const char *arg3;

    if (!PyArg_ParseTuple(args, "sss:call", &arg1, &arg2, &arg3)) {
        return NULL;
    }
    result = PyUnicode_FromFormat(
        "Returning -- value: [%d] arg1: [%s] arg2: [%s] arg3: [%s]\n",
        obj->obj_UnderlyingDatatypePtr->size,
        arg1, arg2, arg3);
    return result;
}
```

```
/* Iterators */
getiterfunc tp_iter;
iternextfunc tp_iternext;
```

These functions provide support for the iterator protocol. Both handlers take exactly one parameter, the instance for which they are being called, and return a new reference. In the case of an error, they should set an exception and return *NULL*. `tp_iter` corresponds to the Python `__iter__()` method, while `tp_iternext` corresponds to the Python `__next__()` method.

Any *iterable* object must implement the `tp_iter` handler, which must return an *iterator* object. Here the same guidelines apply as for Python classes:

- For collections (such as lists and tuples) which can support multiple independent iterators, a new iterator should be created and returned by each call to `tp_iter`.
- Objects which can only be iterated over once (usually due to side effects of iteration, such as file objects) can implement `tp_iter` by returning a new reference to themselves – and should also therefore implement the `tp_iternext` handler.

Any *iterator* object should implement both `tp_iter` and `tp_iternext`. An iterator's `tp_iter` handler should return a new reference to the iterator. Its `tp_iternext` handler should return a new reference to the next object in the iteration, if there is one. If the iteration has reached the end, `tp_iternext` may return *NULL* without setting an exception, or it may set `StopIteration` *in addition* to returning *NULL*; avoiding the exception can yield slightly better performance. If an actual error occurs, `tp_iternext` should always set an exception and return *NULL*.

## 2.3.6 Weak Reference Support

One of the goals of Python's weak reference implementation is to allow any type to participate in the weak reference mechanism without incurring the overhead on performance-critical objects (such as numbers).

参见:

Documentation for the `weakref` module.

For an object to be weakly referencable, the extension type must do two things:

1. Include a `PyObject*` field in the C object structure dedicated to the weak reference mechanism. The object's constructor should leave it *NULL* (which is automatic when using the default `tp_alloc`).
2. Set the `tp_weaklistoffset` type member to the offset of the aforementioned field in the C object structure, so that the interpreter knows how to access and modify that field.

Concretely, here is how a trivial object structure would be augmented with the required field:

```
typedef struct {
    PyObject_HEAD
    PyObject *weakreflist; /* List of weak references */
} TrivialObject;
```

And the corresponding member in the statically-declared type object:

```
static PyTypeObject TrivialType = {
    PyVarObject_HEAD_INIT(NULL, 0)
    /* ... other members omitted for brevity ... */
    .tp_weaklistoffset = offsetof(TrivialObject, weakreflist),
};
```

The only further addition is that `tp_dealloc` needs to clear any weak references (by calling `PyObject_ClearWeakRefs()`) if the field is non-*NULL*:

```
static void
Trivial_dealloc(TrivialObject *self)
{
    /* Clear weakrefs first before calling any destructors */
    if (self->weakreflist != NULL)
        PyObject_ClearWeakRefs((PyObject *) self);
    /* ... remainder of destruction code omitted for brevity ... */
    Py_TYPE(self)->tp_free((PyObject *) self);
}
```

## 2.3.7 更多建议

In order to learn how to implement any specific method for your new data type, get the *CPython* source code. Go to the `Objects` directory, then search the C source files for `tp_` plus the function you want (for example, `tp_richcompare`). You will find examples of the function you want to implement.

When you need to verify that an object is a concrete instance of the type you are implementing, use the `PyObject_TypeCheck()` function. A sample of its use might be something like the following:

```
if (!PyObject_TypeCheck(some_object, &MyType)) {
    PyErr_SetString(PyExc_TypeError, "arg #1 not a mything");
    return NULL;
}
```

参见:

下载 CPython 源代码版本。 <https://www.python.org/downloads/source/>

GitHub 上开发 CPython 源代码的 CPython 项目。 <https://github.com/python/cpython>

## 2.4 构建 C/C++ 扩展

一个 CPython 的 C 扩展是一个共享库 (例如一个 Linux 上的 `.so` , 或者 Windows 上的 `.pyd`), 其会导出一个初始化函数。

为了可导入, 共享库必须在 `PYTHONPATH` 中有效, 且必须命名遵循模块名字, 通过适当的扩展。当使用 `distutils` 时, 会自动生成正确的文件名。

初始化函数的声明如下:

`PyObject* PyInit_modulename (void)`

该函数返回完整初始化过的模块, 或一个 `PyModuleDef` 实例。查看 `initializing-modules` 了解更多细节。

对于仅有 ASCII 编码的模块名, 函数必须是 `PyInit_<modulename>` , 将 `<modulename>` 替换为模块的名字。当使用 `multi-phase-initialization` 时, 允许使用非 ASCII 编码的模块名。此时初始化函数的名字是 `PyInitU_<modulename>` , 而 `<modulename>` 需要用 Python 的 *punycode* 编码, 连字号需替换为下划线。在 Python 里:

```
def initfunc_name(name):
    try:
        suffix = b'_' + name.encode('ascii')
    except UnicodeEncodeError:
        suffix = b'U_' + name.encode('punycode').replace(b'-', b'_')
    return b'PyInit' + suffix
```

可以在一个动态库里导出多个模块，通过定义多个初始化函数。而导入他们需要符号链接或自定义导入器，因为缺省时只有对应了文件名的函数才会被发现。查看“一个库里的多模块”章节，在 [PEP 489](#) 了解更多细节。

## 2.4.1 使用 distutils 构建 C 和 C++ 扩展

扩展模块可以用 distutils 来构建，这是 Python 自带的。distutils 也支持创建二进制包，用户无需编译器而 distutils 就能安装扩展。

一个 distutils 包包含了一个驱动脚本 setup.py。这是个纯 Python 文件，大多数时候也很简单，看起来如下：

```
from distutils.core import setup, Extension

module1 = Extension('demo',
                    sources = ['demo.c'])

setup (name = 'PackageName',
       version = '1.0',
       description = 'This is a demo package',
       ext_modules = [module1])
```

通过文件 setup.py，和文件 demo.c，运行如下

```
python setup.py build
```

这会编译 demo.c，然后产生一个扩展模块叫做 demo 在目录 build 里。依赖于系统，模块文件会放在某个子目录形如 build/lib.system，名字可能是 demo.so 或 demo.pyd。

在文件 setup.py 里，所有动作的入口通过 setup 函数。该函数可以接受可变数量个关键字参数，上面的例子只使用了一个子集。特别需要注意的例子指定了构建包的元信息，以及指定了包内容。通常一个包会包括多个模块，就像 Python 的源码模块、文档、子包等。请参数 distutils 的文档，在 distutils-index 来了解更多 distutils 的特性；本章节只解释构建扩展模块的部分。

通常预计算参数给 setup()，想要更好的结构化驱动脚本。有如如上例子函数 setup() 的 ext\_modules 参数是一列扩展模块，每个是一个 Extension 类的实例。例子中的实例定义了扩展命名为 demo，从单一源码文件构建 demo.c。

更多时候，构建一个扩展会复杂的多，需要额外的预处理器定义和库。如下例子展示了这些。

```
from distutils.core import setup, Extension

module1 = Extension('demo',
                    define_macros = [('MAJOR_VERSION', '1'),
                                     ('MINOR_VERSION', '0')],
                    include_dirs = ['/usr/local/include'],
                    libraries = ['tcl83'],
                    library_dirs = ['/usr/local/lib'],
                    sources = ['demo.c'])

setup (name = 'PackageName',
       version = '1.0',
       description = 'This is a demo package',
       author = 'Martin v. Loewis',
       author_email = 'martin@v.loewis.de',
       url = 'https://docs.python.org/extending/building',
       long_description = '''
This is really just a demo package.
```

(下页继续)

(续上页)

```
'''
    ext_modules = [module1])
```

例子中函数 `setup()` 在调用时额外传递了元信息，是推荐发布包构建时的内容。对于这个扩展，其指定了预处理器定义，`include` 目录，库目录，库。依赖于编译器，`distutils` 还会用其他方式传递信息给编译器。例如在 Unix 上，结果是如下编译命令

```
gcc -DNDEBUG -g -O3 -Wall -Wstrict-prototypes -fPIC -DMAJOR_VERSION=1 -DMINOR_
↪VERSION=0 -I/usr/local/include -I/usr/local/include/python2.2 -c demo.c -o build/
↪temp.linux-i686-2.2/demo.o

gcc -shared build/temp.linux-i686-2.2/demo.o -L/usr/local/lib -ltcl83 -o build/lib.
↪linux-i686-2.2/demo.so
```

这些行代码仅用于展示目的；`distutils` 用户应该相信 `distutils` 能正确调用。

## 2.4.2 发布你的扩展模块

When an extension has been successfully built, there are three ways to use it.

最终用户通常想要安装模块，可以这么运行

```
python setup.py install
```

模块维护者应该制作源码包；要实现可以运行

```
python setup.py sdist
```

有些情况下，需要在源码发布包里包含额外的文件；这通过 `MANIFEST.in` 文件实现，查看 `manifest` 了解细节。

If the source distribution has been built successfully, maintainers can also create binary distributions. Depending on the platform, one of the following commands can be used to do so.

```
python setup.py bdist_wininst
python setup.py bdist_rpm
python setup.py bdist_dumb
```

## 2.5 在 Windows 平台编译 C 和 C++ 扩展

This chapter briefly explains how to create a Windows extension module for Python using Microsoft Visual C++, and follows with more detailed background information on how it works. The explanatory material is useful for both the Windows programmer learning to build Python extensions and the Unix programmer interested in producing software which can be successfully built on both Unix and Windows.

Module authors are encouraged to use the `distutils` approach for building extension modules, instead of the one described in this section. You will still need the C compiler that was used to build Python; typically Microsoft Visual C++.

**注解：** This chapter mentions a number of filenames that include an encoded Python version number. These filenames are represented with the version number shown as `XY`; in practice, `'X'` will be the major version number and `'Y'` will be the minor version number of the Python release you're working with. For example, if you are using Python 2.2.1, `XY` will actually be `22`.

## 2.5.1 A Cookbook Approach

There are two approaches to building extension modules on Windows, just as there are on Unix: use the `distutils` package to control the build process, or do things manually. The `distutils` approach works well for most extensions; documentation on using `distutils` to build and package extension modules is available in `distutils-index`. If you find you really need to do things manually, it may be instructive to study the project file for the `winsound` standard library module.

## 2.5.2 Differences Between Unix and Windows

Unix and Windows use completely different paradigms for run-time loading of code. Before you try to build a module that can be dynamically loaded, be aware of how your system works.

In Unix, a shared object (`.so`) file contains code to be used by the program, and also the names of functions and data that it expects to find in the program. When the file is joined to the program, all references to those functions and data in the file's code are changed to point to the actual locations in the program where the functions and data are placed in memory. This is basically a link operation.

In Windows, a dynamic-link library (`.dll`) file has no dangling references. Instead, an access to functions or data goes through a lookup table. So the DLL code does not have to be fixed up at runtime to refer to the program's memory; instead, the code already uses the DLL's lookup table, and the lookup table is modified at runtime to point to the functions and data.

In Unix, there is only one type of library file (`.a`) which contains code from several object files (`.o`). During the link step to create a shared object file (`.so`), the linker may find that it doesn't know where an identifier is defined. The linker will look for it in the object files in the libraries; if it finds it, it will include all the code from that object file.

In Windows, there are two types of library, a static library and an import library (both called `.lib`). A static library is like a Unix `.a` file; it contains code to be included as necessary. An import library is basically used only to reassure the linker that a certain identifier is legal, and will be present in the program when the DLL is loaded. So the linker uses the information from the import library to build the lookup table for using identifiers that are not included in the DLL. When an application or a DLL is linked, an import library may be generated, which will need to be used for all future DLLs that depend on the symbols in the application or DLL.

Suppose you are building two dynamic-load modules, B and C, which should share another block of code A. On Unix, you would *not* pass `A.a` to the linker for `B.so` and `C.so`; that would cause it to be included twice, so that B and C would each have their own copy. In Windows, building `A.dll` will also build `A.lib`. You *do* pass `A.lib` to the linker for B and C. `A.lib` does not contain code; it just contains information which will be used at runtime to access A's code.

In Windows, using an import library is sort of like using `import spam`; it gives you access to `spam`'s names, but does not create a separate copy. On Unix, linking with a library is more like `from spam import *`; it does create a separate copy.

## 2.5.3 Using DLLs in Practice

Windows Python is built in Microsoft Visual C++; using other compilers may or may not work (though Borland seems to). The rest of this section is MSVC++ specific.

When creating DLLs in Windows, you must pass `pythonXY.lib` to the linker. To build two DLLs, `spam` and `ni` (which uses C functions found in `spam`), you could use these commands:

```
cl /LD /I/python/include spam.c ../libs/pythonXY.lib
cl /LD /I/python/include ni.c spam.lib ../libs/pythonXY.lib
```

The first command created three files: `spam.obj`, `spam.dll` and `spam.lib`. `Spam.dll` does not contain any Python functions (such as `PyArg_ParseTuple()`), but it does know how to find the Python code thanks to `pythonXY.lib`.

The second command created `ni.dll` (and `.obj` and `.lib`), which knows how to find the necessary functions from `spam`, and also from the Python executable.

Not every identifier is exported to the lookup table. If you want any other modules (including Python) to be able to see your identifiers, you have to say `_declspec(dllexport)`, as in `void _declspec(dllexport) initspam(void)` or `PyObject _declspec(dllexport) *NiGetSpamData(void)`.

Developer Studio will throw in a lot of import libraries that you do not really need, adding about 100K to your executable. To get rid of them, use the Project Settings dialog, Link tab, to specify *ignore default libraries*. Add the correct `msvcrtxx.lib` to the list of libraries.





---

## 在更大的应用程序中嵌入 CPython 运行时

---

有时，不是要创建在 Python 解释器中作为主应用程序运行的扩展，而是希望将 CPython 运行时嵌入到更大的应用程序中。本节介绍了成功完成此操作所涉及的一些细节。

### 3.1 Embedding Python in Another Application

The previous chapters discussed how to extend Python, that is, how to extend the functionality of Python by attaching a library of C functions to it. It is also possible to do it the other way around: enrich your C/C++ application by embedding Python in it. Embedding provides your application with the ability to implement some of the functionality of your application in Python rather than C or C++. This can be used for many purposes; one example would be to allow users to tailor the application to their needs by writing some scripts in Python. You can also use it yourself if some of the functionality can be written in Python more easily.

Embedding Python is similar to extending it, but not quite. The difference is that when you extend Python, the main program of the application is still the Python interpreter, while if you embed Python, the main program may have nothing to do with Python — instead, some parts of the application occasionally call the Python interpreter to run some Python code.

So if you are embedding Python, you are providing your own main program. One of the things this main program has to do is initialize the Python interpreter. At the very least, you have to call the function `Py_Initialize()`. There are optional calls to pass command line arguments to Python. Then later you can call the interpreter from any part of the application.

There are several different ways to call the interpreter: you can pass a string containing Python statements to `PyRun_SimpleString()`, or you can pass a stdio file pointer and a file name (for identification in error messages only) to `PyRun_SimpleFile()`. You can also call the lower-level operations described in the previous chapters to construct and use Python objects.

**参见：**

**c-api-index** The details of Python's C interface are given in this manual. A great deal of necessary information can be found here.

### 3.1.1 Very High Level Embedding

The simplest form of embedding Python is the use of the very high level interface. This interface is intended to execute a Python script without needing to interact with the application directly. This can for example be used to perform some operation on a file.

```
#define PY_SSIZE_T_CLEAN
#include <Python.h>

int
main(int argc, char *argv[])
{
    wchar_t *program = Py_DecodeLocale(argv[0], NULL);
    if (program == NULL) {
        fprintf(stderr, "Fatal error: cannot decode argv[0]\n");
        exit(1);
    }
    Py_SetProgramName(program); /* optional but recommended */
    Py_Initialize();
    PyRun_SimpleString("from time import time,ctime\n"
                      "print('Today is', ctime(time()))\n");
    if (Py_FinalizeEx() < 0) {
        exit(120);
    }
    PyMem_RawFree(program);
    return 0;
}
```

The `Py_SetProgramName()` function should be called before `Py_Initialize()` to inform the interpreter about paths to Python run-time libraries. Next, the Python interpreter is initialized with `Py_Initialize()`, followed by the execution of a hard-coded Python script that prints the date and time. Afterwards, the `Py_FinalizeEx()` call shuts the interpreter down, followed by the end of the program. In a real program, you may want to get the Python script from another source, perhaps a text-editor routine, a file, or a database. Getting the Python code from a file can better be done by using the `PyRun_SimpleFile()` function, which saves you the trouble of allocating memory space and loading the file contents.

### 3.1.2 Beyond Very High Level Embedding: An overview

The high level interface gives you the ability to execute arbitrary pieces of Python code from your application, but exchanging data values is quite cumbersome to say the least. If you want that, you should use lower level calls. At the cost of having to write more C code, you can achieve almost anything.

It should be noted that extending Python and embedding Python is quite the same activity, despite the different intent. Most topics discussed in the previous chapters are still valid. To show this, consider what the extension code from Python to C really does:

1. Convert data values from Python to C,
2. Perform a function call to a C routine using the converted values, and
3. Convert the data values from the call from C to Python.

When embedding Python, the interface code does:

1. Convert data values from C to Python,
2. Perform a function call to a Python interface routine using the converted values, and
3. Convert the data values from the call from Python to C.

As you can see, the data conversion steps are simply swapped to accommodate the different direction of the cross-language transfer. The only difference is the routine that you call between both data conversions. When extending, you call a C routine, when embedding, you call a Python routine.

This chapter will not discuss how to convert data from Python to C and vice versa. Also, proper use of references and dealing with errors is assumed to be understood. Since these aspects do not differ from extending the interpreter, you can refer to earlier chapters for the required information.

### 3.1.3 Pure Embedding

The first program aims to execute a function in a Python script. Like in the section about the very high level interface, the Python interpreter does not directly interact with the application (but that will change in the next section).

The code to run a function defined in a Python script is:

```
#define PY_SSIZE_T_CLEAN
#include <Python.h>

int
main(int argc, char *argv[])
{
    PyObject *pName, *pModule, *pFunc;
    PyObject *pArgs, *pValue;
    int i;

    if (argc < 3) {
        fprintf(stderr, "Usage: call pythonfile funcname [args]\n");
        return 1;
    }

    Py_Initialize();
    pName = PyUnicode_DecodeFSDefault(argv[1]);
    /* Error checking of pName left out */

    pModule = PyImport_Import(pName);
    Py_DECREF(pName);

    if (pModule != NULL) {
        pFunc = PyObject_GetAttrString(pModule, argv[2]);
        /* pFunc is a new reference */

        if (pFunc && PyCallable_Check(pFunc)) {
            pArgs = PyTuple_New(argc - 3);
            for (i = 0; i < argc - 3; ++i) {
                pValue = PyLong_FromLong(atoi(argv[i + 3]));
                if (!pValue) {
                    Py_DECREF(pArgs);
                    Py_DECREF(pModule);
                    fprintf(stderr, "Cannot convert argument\n");
                    return 1;
                }
                /* pValue reference stolen here: */
                PyTuple_SetItem(pArgs, i, pValue);
            }
            pValue = PyObject_CallObject(pFunc, pArgs);
            Py_DECREF(pArgs);
            if (pValue != NULL) {
```

(下页继续)

(续上页)

```

        printf("Result of call: %ld\n", PyLong_AsLong(pValue));
        Py_DECREF(pValue);
    }
    else {
        Py_DECREF(pFunc);
        Py_DECREF(pModule);
        PyErr_Print();
        fprintf(stderr, "Call failed\n");
        return 1;
    }
}
else {
    if (PyErr_Occurred())
        PyErr_Print();
    fprintf(stderr, "Cannot find function \"%s\"\n", argv[2]);
}
Py_XDECREF(pFunc);
Py_DECREF(pModule);
}
else {
    PyErr_Print();
    fprintf(stderr, "Failed to load \"%s\"\n", argv[1]);
    return 1;
}
if (Py_FinalizeEx() < 0) {
    return 120;
}
return 0;
}

```

This code loads a Python script using `argv[1]`, and calls the function named in `argv[2]`. Its integer arguments are the other values of the `argv` array. If you *compile and link* this program (let's call the finished executable **call**), and use it to execute a Python script, such as:

```

def multiply(a,b):
    print("Will compute", a, "times", b)
    c = 0
    for i in range(0, a):
        c = c + b
    return c

```

then the result should be:

```

$ call multiply multiply 3 2
Will compute 3 times 2
Result of call: 6

```

Although the program is quite large for its functionality, most of the code is for data conversion between Python and C, and for error reporting. The interesting part with respect to embedding Python starts with

```

Py_Initialize();
pName = PyUnicode_DecodeFSDefault(argv[1]);
/* Error checking of pName left out */
pModule = PyImport_Import(pName);

```

After initializing the interpreter, the script is loaded using `PyImport_Import()`. This routine needs a Python string as its argument, which is constructed using the `PyUnicode_FromString()` data conversion routine.

```
pFunc = PyObject_GetAttrString(pModule, argv[2]);
/* pFunc is a new reference */

if (pFunc && PyCallable_Check(pFunc)) {
    ...
}
Py_XDECREF(pFunc);
```

Once the script is loaded, the name we're looking for is retrieved using `PyObject_GetAttrString()`. If the name exists, and the object returned is callable, you can safely assume that it is a function. The program then proceeds by constructing a tuple of arguments as normal. The call to the Python function is then made with:

```
pValue = PyObject_CallObject(pFunc, pArgs);
```

Upon return of the function, `pValue` is either `NULL` or it contains a reference to the return value of the function. Be sure to release the reference after examining the value.

### 3.1.4 Extending Embedded Python

Until now, the embedded Python interpreter had no access to functionality from the application itself. The Python API allows this by extending the embedded interpreter. That is, the embedded interpreter gets extended with routines provided by the application. While it sounds complex, it is not so bad. Simply forget for a while that the application starts the Python interpreter. Instead, consider the application to be a set of subroutines, and write some glue code that gives Python access to those routines, just like you would write a normal Python extension. For example:

```
static int numargs=0;

/* Return the number of arguments of the application command line */
static PyObject*
emb_numargs(PyObject *self, PyObject *args)
{
    if(!PyArg_ParseTuple(args, ":numargs"))
        return NULL;
    return PyLong_FromLong(numargs);
}

static PyMethodDef EmbMethods[] = {
    {"numargs", emb_numargs, METH_VARARGS,
     "Return the number of arguments received by the process."},
    {NULL, NULL, 0, NULL}
};

static PyModuleDef EmbModule = {
    PyModuleDef_HEAD_INIT, "emb", NULL, -1, EmbMethods,
    NULL, NULL, NULL, NULL
};

static PyObject*
PyInit_emb(void)
{
    return PyModule_Create(&EmbModule);
}
```

Insert the above code just above the `main()` function. Also, insert the following two statements before the call to `Py_Initialize()`:

```
numargs = argc;
PyImport_AppendInittab("emb", &PyInit_emb);
```

These two lines initialize the `numargs` variable, and make the `emb.numargs()` function accessible to the embedded Python interpreter. With these extensions, the Python script can do things like

```
import emb
print("Number of arguments", emb.numargs())
```

In a real application, the methods will expose an API of the application to Python.

### 3.1.5 Embedding Python in C++

It is also possible to embed Python in a C++ program; precisely how this is done will depend on the details of the C++ system used; in general you will need to write the main program in C++, and use the C++ compiler to compile and link your program. There is no need to recompile Python itself using C++.

### 3.1.6 Compiling and Linking under Unix-like systems

It is not necessarily trivial to find the right flags to pass to your compiler (and linker) in order to embed the Python interpreter into your application, particularly because Python needs to load library modules implemented as C dynamic extensions (`.so` files) linked against it.

To find out the required compiler and linker flags, you can execute the `pythonX.Y-config` script which is generated as part of the installation process (a `python3-config` script may also be available). This script has several options, of which the following will be directly useful to you:

- `pythonX.Y-config --cflags` will give you the recommended flags when compiling:

```
$ /opt/bin/python3.4-config --cflags
-I/opt/include/python3.4m -I/opt/include/python3.4m -DNDEBUG -g -fwrapv -O3 -Wall
↪-Wstrict-prototypes
```

- `pythonX.Y-config --ldflags` will give you the recommended flags when linking:

```
$ /opt/bin/python3.4-config --ldflags
-L/opt/lib/python3.4/config-3.4m -lpthread -ldl -lutil -lm -lpython3.4m -Xlinker -
↪export-dynamic
```

**注解:** To avoid confusion between several Python installations (and especially between the system Python and your own compiled Python), it is recommended that you use the absolute path to `pythonX.Y-config`, as in the above example.

If this procedure doesn't work for you (it is not guaranteed to work for all Unix-like platforms; however, we welcome bug reports) you will have to read your system's documentation about dynamic linking and/or examine Python's Makefile (use `sysconfig.get_makefile_filename()` to find its location) and compilation options. In this case, the `sysconfig` module is a useful tool to programmatically extract the configuration values that you will want to combine together. For example:

```
>>> import sysconfig
>>> sysconfig.get_config_var('LIBS')
'-lpthread -ldl -lutil'
>>> sysconfig.get_config_var('LINKFORSHARED')
'-Xlinker -export-dynamic'
```

## 术语对照表

>>> 交互式终端中默认的 Python 提示符。往往会显示于能以交互方式在解释器里执行的样例代码之前。

... Can refer to:

- 交互式终端中输入特殊代码行时默认的 Python 提示符，包括：缩进的代码块，成对的分隔符之内（圆括号、方括号、花括号或三重引号），或是指定一个装饰器之后。
- The Ellipsis built-in constant.

**2to3** 一个将 Python 2.x 代码转换为 Python 3.x 代码的工具，能够处理大部分通过解析源码并遍历解析树可检测到的不兼容问题。

2to3 包含在标准库中，模块名为 `lib2to3`；并提供一个独立入口点 `Tools/scripts/2to3`。参见 `2to3-reference`。

**abstract base class – 抽象基类** 抽象基类简称 ABC，是对 *duck-typing* 的补充，它提供了一种定义接口的新方式，相比之下其他技巧例如 `hasattr()` 显得过于笨拙或有微妙错误（例如使用 魔术方法）。ABC 引入了虚拟子类，这种类并非继承自其他类，但却仍能被 `isinstance()` 和 `issubclass()` 所认可；详见 `abc` 模块文档。Python 自带许多内置的 ABC 用于实现数据结构（在 `collections.abc` 模块中）、数字（在 `numbers` 模块中）、流（在 `io` 模块中）、导入查找器和加载器（在 `importlib.abc` 模块中）。你可以使用 `abc` 模块来创建自己的 ABC。

**annotation – 标注** 关联到某个变量、类属性、函数形参或返回值的标签，被约定作为 *type hint* 来使用。

局部变量的标注在运行时不可访问，但全局变量、类属性和函数的标注会分别存放模块、类和函数的 `__annotations__` 特殊属性中。

参见 *variable annotation*、*function annotation*、**PEP 484** 和 **PEP 526**，对此功能均有介绍。

**argument – 参数** 在调用函数时传给 *function*（或 *method*）的值。参数分为两种：

- 关键字参数：在函数调用中前面带有标识符（例如 `name=`）或者作为包含在前面带有 `**` 的字典里的值传入。举例来说，3 和 5 在以下对 `complex()` 的调用中均属于关键字参数：

```
complex(real=3, imag=5)
complex(**{'real': 3, 'imag': 5})
```

- 位置参数: 不属于关键字参数的参数。位置参数可出现于参数列表的开头以及/或者作为前面带有 \* 的 *iterable* 里的元素被传入。举例来说, 3 和 5 在以下调用中均属于位置参数:

```
complex(3, 5)
complex(*(3, 5))
```

参数会被赋值给函数体中对应的局部变量。有关赋值规则参见 *calls* 一节。根据语法, 任何表达式都可用来表示一个参数; 最终算出的值会被赋给对应的局部变量。

另参见 *parameter* 术语表条目, 常见问题中 参数与形参的区别以及 **PEP 362**。

**asynchronous context manager – 异步上下文管理器** 此种对象通过定义 `__aenter__()` 和 `__aexit__()` 方法来对 `async with` 语句中的环境进行控制。由 **PEP 492** 引入。

**asynchronous generator – 异步生成器** 返回值为 *asynchronous generator iterator* 的函数。它与使用 `async def` 定义的协程函数很相似, 不同之处在于它包含 `yield` 表达式以产生一系列可在 `async for` 循环中使用的值。

此术语通常是指异步生成器函数, 但在某些情况下则可能是指 异步生成器迭代器。如果需要清楚表达具体含义, 请使用全称以避免歧义。

一个异步生成器函数可能包含 `await` 表达式或者 `async for` 以及 `async with` 语句。

**asynchronous generator iterator – 异步生成器迭代器** *asynchronous generator* 函数所创建的对象。

此对象属于 *asynchronous iterator*, 当使用 `__anext__()` 方法调用时会返回一个可等待对象来执行异步生成器函数的代码直到下一个 `yield` 表达式。

每个 `yield` 会临时暂停处理, 记住当前位置执行状态 (包括局部变量和挂起的 `try` 语句)。当该 异步生成器迭代器与其他 `__anext__()` 返回的可等待对象有效恢复时, 它会从离开位置继续执行。参见 **PEP 492** 和 **PEP 525**。

**asynchronous iterable – 异步可迭代对象** 可在 `async for` 语句中被使用的对象。必须通过它的 `__aiter__()` 方法返回一个 *asynchronous iterator*。由 **PEP 492** 引入。

**asynchronous iterator – 异步迭代器** 实现了 `__aiter__()` 和 `__anext__()` 方法的对象。`__anext__` 必须返回一个 *awaitable* 对象。`async for` 会处理异步迭代器的 `__anext__()` 方法所返回的可等待对象, 直到其引发一个 `StopAsyncIteration` 异常。由 **PEP 492** 引入。

**attribute – 属性** 关联到一个对象的值, 可以使用点号表达式通过其名称来引用。例如, 如果一个对象 *o* 具有一个属性 *a*, 就可以用 *o.a* 来引用它。

**awaitable – 可等待对象** 能在 `await` 表达式中使用的对象。可以是 *coroutine* 或是具有 `__await__()` 方法的对象。参见 **PEP 492**。

**BDFL** “终身仁慈独裁者”的英文缩写, 即 *Guido van Rossum*, Python 的创造者。

**binary file – 二进制文件** *file object* 能够读写字节类对象。二进制文件的例子包括以二进制模式 ('rb', 'wb' 或 'rb+') 打开的文件、`sys.stdin.buffer`、`sys.stdout.buffer` 以及 `io.BytesIO` 和 `gzip.GzipFile` 的实例。

另请参见 *text file* 了解能够读写 `str` 对象的文件对象。

**bytes-like object – 字节类对象** 支持 *bufferobjects* 并且能导出 *C-contiguous* 缓冲的对象。这包括所有 `bytes`、`bytearray` 和 `array.array` 对象, 以及许多普通 *memoryview* 对象。字节类对象可在多种二进制数据操作中使用; 这些操作包括压缩、保存为二进制文件以及通过套接字发送等。

某些操作需要可变的二进制数据。这种对象在文档中常被称为“可读写字节类对象”。可变缓冲对象的例子包括 `bytearray` 以及 `bytearray` 的 *memoryview*。其他操作要求二进制数据存放于不可变对象 (“只读字节类对象”); 这种对象的例子包括 `bytes` 以及 `bytes` 对象的 *memoryview*。

**bytecode – 字节码** Python 源代码会被编译为字节码, 即 CPython 解释器中表示 Python 程序的内部代码。字节码还会缓存在 `.pyc` 文件中, 这样第二次执行同一文件时速度更快 (可以免去将源码重新编译为字



节码)。这种“中间语言”运行在根据字节码执行相应机器码的 *virtual machine* 之上。请注意不同 Python 虚拟机上的字节码不一定通用，也不一定能在不同 Python 版本上兼容。

字节码指令列表可以在 `dis` 模块的文档中查看。

**class – 类** 用来创建用户定义对象的模板。类定义通常包含对该类的实例进行操作的方法定义。

**class variable – 类变量** 在类中定义的变量，并且仅限在类的层级上修改（而不是在类的实例中修改）。

**coercion – 强制类型转换** 在包含两个相同类型参数的操作中，一种类型的实例隐式地转换为另一种类型。例如，`int(3.15)` 是将原浮点数转换为整型数 3，但在 `3+4.5` 中，参数的类型不一致（一个是 `int`，一个是 `float`），两者必须转换为相同类型才能相加，否则将引发 `TypeError`。如果没有强制类型转换机制，程序员必须将所有可兼容参数归一化为相同类型，例如要写成 `float(3)+4.5` 而不是 `3+4.5`。

**complex number – 复数** 对普通实数系统的扩展，其中所有数字都被表示为一个实部和一个虚部的和。虚数是虚数单位（-1 的平方根）的实倍数，通常在数学中写为 `i`，在工程学中写为 `j`。Python 内置了对复数的支持，采用工程学标记方式；虚部带有一个 `j` 后缀，例如 `3+1j`。如果需要 `math` 模块内对象的对应复数版本，请使用 `cmath`，复数的使用是一个比较高级的数学特性。如果你感觉没有必要，忽略它们也几乎不会有任何问题。

**context manager – 上下文管理器** 在 `with` 语句中使用，通过定义 `__enter__()` 和 `__exit__()` 方法来控制环境状态的对象。参见 [PEP 343](#)。

**context variable – 上下文变量** 一种根据其所属的上下文可以具有不同的值的变量。这类似于在线程局部存储中每个执行线程可以具有不同的变量值。不过，对于上下文变量来说，一个执行线程中可能会有多个上下文，而上下文变量的主要用途是对并发异步任务中变量进行追踪。参见 `contextvars`。

**contiguous – 连续** 一个缓冲如果是 C 连续或 Fortran 连续就会被认为是连续的。零维缓冲是 C 和 Fortran 连续的。在一维数组中，所有条目必须在内存中彼此相邻地排列，采用从零开始的递增索引顺序。在多维 C-连续数组中，当按内存地址排列时用最后一个索引访问条目时速度最快。但是在 Fortran 连续数组中则是用第一个索引最快。

**coroutine – 协程** 协程是子例程的更一般形式。子例程可以在某一点进入并在另一点退出。协程则可以在许多不同的点上进入、退出和恢复。它们可通过 `async def` 语句来实现。参见 [PEP 492](#)。

**coroutine function – 协程函数** 返回一个 *coroutine* 对象的函数。协程函数可通过 `async def` 语句来定义，并可能包含 `await`、`async for` 和 `async with` 关键字。这些特性是由 [PEP 492](#) 引入的。

**CPython** Python 编程语言的规范实现，在 [python.org](#) 上发布。“CPython”一词用于在必要时将此实现与其他实现例如 Jython 或 IronPython 相区别。

**decorator – 装饰器** 返回值为另一个函数的函数，通常使用 `@wrapper` 语法形式来进行函数变换。装饰器的常见例子包括 `classmethod()` 和 `staticmethod()`。

装饰器语法只是一种语法糖，以下两个函数定义在语义上完全等价：

```
def f(...):
    ...
f = staticmethod(f)

@staticmethod
def f(...):
    ...
```

同样的概念也适用于类，但通常较少这样使用。有关装饰器的详情可参见 [函数定义](#) 和 [类定义](#) 的文档。

**descriptor – 描述器** 任何定义了 `__get__()`、`__set__()` 或 `__delete__()` 方法的对象。当一个类属性为描述器时，它的特殊绑定行为就会在属性查找时被触发。通常情况下，使用 `a.b` 来获取、设置或删除一个属性时会在 `a` 的类字典中查找名称为 `b` 的对象，但如果 `b` 是一个描述器，则会调用对应的描述器方法。理解描述器的概念是更深层次理解 Python 的关键，因为这是许多重要特性的基础，包括函数、方法、属性、类方法、静态方法以及对超类的引用等等。

有关描述符的方法的详情可参看 [descriptors](#)。

**dictionary – 字典** 一个关联数组，其中的任意键都映射到相应的值。键可以是任何具有 `__hash__()` 和 `__eq__()` 方法的对象。在 Perl 语言中称为 `hash`。

**dictionary view – 字典视图** 从 `dict.keys()`、`dict.values()` 和 `dict.items()` 返回的对象被称为字典视图。它们提供了字典条目的一个动态视图，这意味着当字典改变时，视图也会相应改变。要将字典视图强制转换为真正的列表，可使用 `list(dictview)`。参见 [dict-views](#)。

**docstring – 文档字符串** 作为类、函数或模块之内的第一个表达式出现的字符串字面值。它在代码执行时会被忽略，但会被解释器识别并放入所在类、函数或模块的 `__doc__` 属性中。由于它可用于代码内省，因此是对象存放文档的规范位置。

**duck-typing – 鸭子类型** 指一种编程风格，它并不依靠查找对象类型来确定其是否具有正确的接口，而是直接调用或使用其方法或属性（“看起来像鸭子，叫起来也像鸭子，那么肯定就是鸭子。”）由于强调接口而非特定类型，设计良好的代码可通过允许多态替代来提升灵活性。鸭子类型避免使用 `type()` 或 `isinstance()` 检测。（但要注意鸭子类型可以使用 [抽象基类](#) 作为补充。）而往往会采用 `hasattr()` 检测或是 [EAFP](#) 编程。

**EAFP “求原谅比求许可更容易”** 的英文缩写。这种 Python 常用代码编写风格会假定所需的键或属性存在，并在假定错误时捕获异常。这种简洁快速风格的特点就是大量运用 `try` 和 `except` 语句。于其相对的则是所谓 [LBYL](#) 风格，常见于 C 等许多其他语言。

**expression – 表达式** 可以求出某个值的语法单元。换句话说，一个表达式就是表达元素例如字面值、名称、属性访问、运算符或函数调用的汇总，它们最终都会返回一个值。与许多其他语言不同，并非所有语言构件都是表达式。还存在不能被用作表达式的 [statement](#)，例如 `while`。赋值也是属于语句而非表达式。

**extension module – 扩展模块** 以 C 或 C++ 编写的模块，使用 Python 的 C API 来与语言核心以及用户代码进行交互。

**f-string – f-字符串** 带有 'f' 或 'F' 前缀的字符串字面值通常被称为“f-字符串”即 格式化字符串字面值的简写。参见 [PEP 498](#)。

**file object – 文件对象** 对外提供面向文件 API 以使用下层资源的对象（带有 `read()` 或 `write()` 这样的方法）。根据其创建方式的不同，文件对象可以处理对真实磁盘文件，对其他类型存储，或是对通讯设备的访问（例如标准输入/输出、内存缓冲区、套接字、管道等等）。文件对象也被称为 文件类对象或 流。

实际上共有三种类别的文件对象：原始 [二进制文件](#)，缓冲 [二进制文件](#) 以及 [文本文件](#)。它们的接口定义均在 `io` 模块中。创建文件对象的规范方式是使用 `open()` 函数。

**file-like object – 文件类对象** [file object](#) 的同义词。

**finder – 查找器** 一种会尝试查找被导入模块的 [loader](#) 的对象。

从 Python 3.3 起存在两种类型的查找器：[元路径查找器](#) 配合 `sys.meta_path` 使用，以及 [path entry finders](#) 配合 `sys.path_hooks` 使用。

更多详情可参见 [PEP 302](#)、[PEP 420](#) 和 [PEP 451](#)。

**floor division – 向下取整除法** 向下舍入到最接近的整数的数学除法。向下取整除法的运算符是 `//`。例如，表达式 `11 // 4` 的计算结果是 2，而与之相反的是浮点数的真正除法返回 2.75。注意 `(-11) // 4` 会返回 -3 因为这是 -2.75 向下舍入得到的结果。见 [PEP 238](#)。

**function – 函数** 可以向调用者返回某个值的一组语句。还可以向其传入零个或多个 [参数](#) 并在函数体执行中被使用。另见 [parameter](#)、[method](#) 和 [function](#) 等节。

**function annotation – 函数标注** 即针对函数形参或返回值的 [annotation](#)。

函数标注通常用于 [类型提示](#)：例如以下函数预期接受两个 `int` 参数并预期返回一个 `int` 值：

```
def sum_two_numbers(a: int, b: int) -> int:
    return a + b
```

函数标注语法的详解见 `function` 一节。

请参看 *variable annotation* 和 **PEP 484** 对此功能的描述。

**\_\_future\_\_** 一种伪模块，可被程序员用来启用与当前解释器不兼容的新语言特性。

通过导入 `__future__` 模块并对其中的变量求值，你可以查看新特性何时首次加入语言以及何时成为默认：

```
>>> import __future__
>>> __future__.division
_Feature((2, 2, 0, 'alpha', 2), (3, 0, 0, 'alpha', 0), 8192)
```

**garbage collection – 垃圾回收** 释放不再被使用的内存空间的过程。Python 是通过引用计数和一个能够检测和打破循环引用的循环垃圾回收器来执行垃圾回收的。可以使用 `gc` 模块来控制垃圾回收器。

**generator – 生成器** 返回一个 *generator iterator* 的函数。它看起来很像普通函数，不同点在于其包含 `yield` 表达式以便产生一系列值供给 `for`-循环使用或是通过 `next()` 函数逐一获取。

通常是指生成器函数，但在某些情况下也可能是指生成器迭代器。如果需要清楚表达具体含义，请使用全称以避免歧义。

**generator iterator – 生成器迭代器** *generator* 函数所创建的对象。

每个 `yield` 会临时暂停处理，记住当前位置执行状态（包括局部变量和挂起的 `try` 语句）。当该生成器迭代器恢复时，它会从离开位置继续执行（这与每次调用都从新开始的普通函数差别很大）。

**generator expression – 生成器表达式** 返回一个迭代器的表达式。它看起来很像普通表达式后面带有定义了一个循环变量、范围的 `for` 子句，以及一个可选的 `if` 子句。以下复合表达式会为外层函数生成一系列值：

```
>>> sum(i*i for i in range(10))           # sum of squares 0, 1, 4, ... 81
285
```

**generic function – 泛型函数** 为不同的类型实现相同操作的多个函数所组成的函数。在调用时会由调度算法来确定应该使用哪个实现。

另请参见 *single dispatch* 术语表条目、`functools.singledispatch()` 装饰器以及 **PEP 443**。

**GIL** 参见 *global interpreter lock*。

**global interpreter lock – 全局解释器锁** CPython 解释器所采用的一种机制，它确保同一时刻只有一个线程在执行 Python *bytecode*。此机制通过设置对象模型（包括 `dict` 等重要内置类型）针对并发访问的隐式安全简化了 CPython 实现。给整个解释器加锁使得解释器多线程运行更方便，其代价则是牺牲了在多处理器上的并行性。

不过，某些标准库或第三方库的扩展模块被设计为在执行计算密集型任务如压缩或哈希时释放 GIL。此外，在执行 I/O 操作时也总是会释放 GIL。

创建一个（以更精细粒度来锁定共享数据的）“自由线程”解释器的努力从未获得成功，因为这会牺牲在普通单处理器情况下的性能。据信克服这种性能问题的措施将导致实现变得更复杂，从而更难以维护。

**hash-based pyc – 基于哈希的 pyc** 使用对应源文件的哈希值而非最后修改时间来确定其有效性的字节码缓存文件。参见 *pyc-invalidation*。

**hashable – 可哈希** 一个对象的哈希值如果在其生命周期内绝不改变，就被称为可哈希（它需要具有 `__hash__()` 方法），并可以同其他对象进行比较（它需要具有 `__eq__()` 方法）。可哈希对象必须具有相同的哈希值比较结果才会相同。

可哈希性使得对象能够作为字典键或集合成员使用，因为这些数据结构要在内部使用哈希值。



大多数 Python 中的不可变内置对象都是可哈希的；可变容器（例如列表或字典）都不可哈希；不可变容器（例如元组和 `frozenset`）仅当它们的元素均为可哈希时才是可哈希的。用户定义类的实例对象默认是可哈希的。它们在比较时一定不相同（除非是与自己比较），它们的哈希值的生成是基于它们的 `id()`。

**IDLE** Python 的 IDE，“集成开发与学习环境”的英文缩写。是 Python 标准发行版附带的基本编程器和解释器环境。

**immutable – 不可变** 具有固定值的对象。不可变对象包括数字、字符串和元组。这样的对象不能被改变。如果必须存储一个不同的值，则必须创建新的对象。它们在需要常量哈希值的地方起着重要作用，例如作为字典中的键。

**import path – 导入路径** 由多个位置（或路径条目）组成的列表，会被模块的 `path based finder` 用来查找导入目标。在导入时，此位置列表通常来自 `sys.path`，但对次级包来说也可能来自上级包的 `__path__` 属性。

**importing – 导入** 令一个模块中的 Python 代码能为另一个模块中的 Python 代码所使用的过程。

**importer – 导入器** 查找并加载模块的对象；此对象既属于 `finder` 又属于 `loader`。

**interactive – 交互** Python 带有一个交互式解释器，即你可以在解释器提示符后输入语句和表达式，立即执行并查看其结果。只需不带参数地启动 `python` 命令（也可以在你的计算机开始菜单中选择相应菜单项）。在测试新想法或检验模块和包的时候用这种方式会非常方便（请记得使用 `help(x)`）。

**interpreted – 解释型** Python 一是种解释型语言，与之相对的是编译型语言，虽然两者的区别由于字节码编译器的存在而会有所模糊。这意味着源文件可以直接运行而不必显式地创建可执行文件再运行。解释型语言通常具有比编译型语言更短的开发/调试周期，但是其程序往往运行得更慢。参见 *interactive*。

**interpreter shutdown – 解释器关闭** 当被要求关闭时，Python 解释器将进入一个特殊运行阶段并逐步释放所有已分配资源，例如模块和各种关键内部结构等。它还会多次调用垃圾回收器。这会触发用户定义析构器或弱引用回调中的代码执行。在关闭阶段执行的代码可能会遇到各种异常，因为其所依赖的资源已不再有效（常见的例子有库模块或警告机制等）。

解释器需要关闭的主要原因有 `__main__` 模块或所运行的脚本已完成执行。

**iterable – 可迭代对象** 能够逐一返回其成员项的对象。可迭代对象的例子包括所有序列类型（例如 `list`、`str` 和 `tuple`）以及某些非序列类型例如 `dict`、文件对象以及定义了 `__iter__()` 方法或是实现了 *Sequence* 语义的 `__getitem__()` 方法的任意自定义类对象。

可迭代对象被可用于 `for` 循环以及许多其他需要一个序列的地方（`zip()`、`map()` ...）。当一个可迭代对象作为参数传给内置函数 `iter()` 时，它会返回该对象的迭代器。这种迭代器适用于对值集合的一次性遍历。在使用可迭代对象时，你通常不需要调用 `iter()` 或者自己处理迭代器对象。`for` 语句会为你自动处理那些操作，创建一个临时的未命名变量用来在循环期间保存迭代器。参见 *iterator*、*sequence* 以及 *generator*。

**iterator – 迭代器** 用来表示一连串数据流的对象。重复调用迭代器的 `__next__()` 方法（或将其传给内置函数 `next()`）将逐个返回流中的项。当没有数据可用时则将引发 `StopIteration` 异常。到这时迭代器对象中的数据项已耗尽，继续调用其 `__next__()` 方法只会再次引发 `StopIteration` 异常。迭代器必须具有 `__iter__()` 方法用来返回该迭代器对象自身，因此迭代器必定也是可迭代对象，可被用于其他可迭代对象适用的大部分场合。一个显著的例外是那些会多次重复访问迭代项的代码。容器对象（例如 `list`）在你每次向其传入 `iter()` 函数或是在 `for` 循环中使用它时都会产生一个新的迭代器。如果在此情况下你尝试用迭代器则会返回在之前迭代过程中被耗尽的同一迭代器对象，使其看起来就像是一个空容器。

更多信息可查看 *typeiter*。

**key function – 键函数** 键函数或称整理函数，是能够返回用于排序或排位的值的可调用对象。例如，`locale.strxfrm()` 可用于生成一个符合特定区域排序约定的排序键。

Python 中有许多工具都允许用键函数来控制元素的排位或分组方式。其中包括 `min()`、`max()`、`sorted()`、`list.sort()`、`heapq.merge()`、`heapq.nsmallest()`、`heapq.nlargest()` 以及

`itertools.groupby()`。

要创建一个键函数有多种方式。例如，`str.lower()` 方法可以用作忽略大小写排序的键函数。另外，键函数也可通过 `lambda` 表达式来创建，例如 `lambda r: (r[0], r[2])`。还有 `operator` 模块提供了三个键函数构造器：`attrgetter()`、`itemgetter()` 和 `methodcaller()`。请查看 [如何排序](#) 一节以获取创建和使用键函数的示例。

**keyword argument – 关键字参数** 参见 [argument](#)。

**lambda** 由一个单独 [expression](#) 构成的匿名内联函数，表达式会在调用时被求值。创建 `lambda` 函数的句法为 `lambda [parameters]: expression`

**LBYL** “先查看后跳跃”的英文缩写。这种代码编写风格会在进行调用或查找之前显式地检查前提条件。此风格与 [EAFP](#) 方式恰成对比，其特点是大量使用 `if` 语句。

在多线程环境中，LBYL 方式会导致“查看”和“跳跃”之间发生条件竞争风险。例如，以下代码 `if key in mapping: return mapping[key]` 可能由于在检查操作之后其他线程从 `mapping` 中移除了 `key` 而出错。这种问题可通过加锁或使用 [EAFP](#) 方式来解决。

**list – 列表** Python 内置的一种 [sequence](#)。虽然名为列表，但更类似于其他语言中的数组而非链接列表，因为访问元素的时间复杂度为  $O(1)$ 。

**list comprehension – 列表推导式** 处理一个序列中的所有或部分元素并返回结果列表的一种紧凑写法。  
`result = ['{:#04x}'.format(x) for x in range(256) if x % 2 == 0]` 将生成一个 0 到 255 范围内的十六进制偶数对应字符串 (0x..) 的列表。其中 `if` 子句是可选的，如果省略则 `range(256)` 中的所有元素都会被处理。

**loader – 加载器** 负责加载模块的对象。它必须定义名为 `load_module()` 的方法。加载器通常由一个 [finder](#) 返回。详情参见 [PEP 302](#)，对于 [abstract base class](#) 可参见 `importlib.abc.Loader`。

**magic method – 魔术方法** [special method](#) 的非正式同义词。

**mapping – 映射** 一种支持任意键查找并实现了 `Mapping` 或 `MutableMapping` 抽象基类中所规定方法的容器对象。此类对象的例子包括 `dict`、`collections.defaultdict`、`collections.OrderedDict` 以及 `collections.Counter`。

**meta path finder – 元路径查找器** `sys.meta_path` 的搜索所返回的 [finder](#)。元路径查找器与 [path entry finders](#) 存在关联但并不相同。

请查看 `importlib.abc.MetaPathFinder` 了解元路径查找器所实现的方法。

**metaclass – 元类** 一种用于创建类的类。类定义包含类名、类字典和基类列表。元类负责接受上述三个参数并创建相应的类。大部分面向对象的编程语言都会提供一个默认实现。Python 的特别之处在于可以创建自定义元类。大部分用户永远不需要这个工具，但当需要出现时，元类可提供强大而优雅的解决方案。它们已被用于记录属性访问日志、添加线程安全性、跟踪对象创建、实现单例，以及其他许多任务。

更多详情参见 [metaclasses](#)。

**method – 方法** 在类内部定义的函数。如果作为该类的实例的一个属性来调用，方法将会获取实例对象作为其第一个 [argument](#) (通常命名为 `self`)。参见 [function](#) 和 [nested scope](#)。

**method resolution order – 方法解析顺序** 方法解析顺序就是在查找成员时搜索全部基类所用的先后顺序。请查看 [Python 2.3 方法解析顺序](#) 了解自 2.3 版起 Python 解析器所用相关算法的详情。

**module – 模块** 此对象是 Python 代码的一种组织单位。各模块具有独立的命名空间，可包含任意 Python 对象。模块可通过 [importing](#) 操作被加载到 Python 中。

另见 [package](#)。

**module spec – 模块规格** 一个命名空间，其中包含用于加载模块的相关导入信息。是 `importlib.machinery.ModuleSpec` 的实例。

**MRO** 参见 [method resolution order](#)。

**mutable – 可变** 可变对象可以在其 `id()` 保持固定的情况下改变其取值。另请参见 *immutable*。

**named tuple – 具名元组** 任何类似元组的类，其中的可索引元素也能使用名称属性来访问。（例如，`time.localtime()` 会返回一个类似元组的对象，其中的 `year` 既可以通过索引访问如 `t[0]` 也可以通过名称属性访问如 `t.tm_year`）。

具名元组可以是一个内置类型例如 `time.struct_time`，也可以通过正规的类定义来创建。一个完备的具名元组还可以通过工厂函数 `collections.namedtuple()` 来创建。后面这种方式会自动提供一些额外特性，例如 `Employee(name='jones', title='programmer')` 这样的自包含文档表示形式。

**namespace – 命名空间** 命名空间是存放变量的场所。命名空间有局部、全局和内置的，还有对象中的嵌套命名空间（在方法之内）。命名空间通过防止命名冲突来支持模块化。例如，函数 `builtins.open` 与 `os.open()` 可通过各自的命名空间来区分。命名空间还通过明确哪个模块实现那个函数来帮助提高可读性和可维护性。例如，`random.seed()` 或 `itertools.islice()` 这种写法明确了这些函数是由 `random` 与 `itertools` 模块分别实现的。

**namespace package – 命名空间包** **PEP 420** 所引入的一种仅被用作子包的容器的 *package*，命名空间包可以没有实体表示物，其描述方式与 *regular package* 不同，因为它们没有 `__init__.py` 文件。

另可参见 *module*。

**nested scope – 嵌套作用域** 在一个定义范围内引用变量的能力。例如，在另一函数之内定义的函数可以引用前者的变量。请注意嵌套作用域默认只对引用有效而对赋值无效。局部变量的读写都受限于最内层作用域。类似的，全局变量的读写则作用于全局命名空间。通过 `nonlocal` 关键字可允许写入外层作用域。

**new-style class – 新式类** 对于目前已被应于所有类对象的类形式的旧称谓。在早先的 Python 版本中，只有新式类能够使用 Python 新增的更灵活特性，例如 `__slots__`、描述符、特征属性、`__getattr__()`、类方法和静态方法等。

**object – 对象** 任何具有状态（属性或值）以及预定义行为（方法）的数据。`object` 也是任何 *new-style class* 的最顶层基类名。

**package – 包** 一种可包含子模块或递归地包含子包的 Python *module*。从技术上说，包是带有 `__path__` 属性的 Python 模块。

另参见 *regular package* 和 *namespace package*。

**parameter – 形参** *function*（或方法）定义中的命名实体，它指定函数可以接受的一个 *argument*（或在某些情况下，多个实参）。有五种形参：

- *positional-or-keyword*：位置或关键字，指定一个可以作为 *位置参数* 传入也可以作为 *关键字参数* 传入的实参。这是默认的形参类型，例如下面的 *foo* 和 *bar*：

```
def func(foo, bar=None): ...
```

- *positional-only*：仅限位置，指定一个只能按位置传入的参数。Python 中没有定义仅限位置形参的语法。但是一些内置函数有仅限位置形参（比如 `abs()`）。
- *keyword-only*：仅限关键字，指定一个只能通过关键字传入的参数。仅限关键字形参可通过在函数定义的形参列表中包含单个可变位置形参或者在多个可变位置形参之前放一个 `*` 来定义，例如下面的 *kw\_only1* 和 *kw\_only2*：

```
def func(arg, *, kw_only1, kw_only2): ...
```

- *var-positional*：可变位置，指定可以提供由一个任意数量的位置参数构成的序列（附加在其他形参已接受的位置参数之后）。这种形参可通过在形参名称前加缀 `*` 来定义，例如下面的 *args*：

```
def func(*args, **kwargs): ...
```



- *var-keyword*: 可变关键字, 指定可以提供任意数量的关键字参数 (附加在其他形参已接受的关键字参数之后)。这种形参可通过在形参名称前加缀 `**` 来定义, 例如上面的 *kwargs*。

形参可以同时指定可选和必选参数, 也可以为某些可选参数指定默认值。

另参见 *argument* 术语表条目、参数与形参的区别中的常见问题、`inspect.Parameter` 类、`function` 一节以及 [PEP 362](#)。

**path entry – 路径入口** *import path* 中的一个单独位置, 会被 *path based finder* 用来查找要导入的模块。

**path entry finder – 路径入口查找器** 任一可调用对象使用 `sys.path_hooks` (即 *path entry hook*) 返回的 *finder*, 此种对象能通过 *path entry* 来定位模块。

请参看 `importlib.abc.PathEntryFinder` 以了解路径入口查找器所实现的各个方法。

**path entry hook – 路径入口钩子** 一种可调用对象, 在知道如何查找特定 *path entry* 中的模块的情况下能够使用 `sys.path_hook` 列表返回一个 *path entry finder*。

**path based finder – 基于路径的查找器** 默认的一种 *元路径查找器*, 可在一个 *import path* 中查找模块。

**path-like object – 路径类对象** 代表一个文件系统路径的对象。类路径对象可以是一个表示路径的 `str` 或者 `bytes` 对象, 还可以是一个实现了 `os.PathLike` 协议的对象。一个支持 `os.PathLike` 协议的对象可通过调用 `os.fspath()` 函数转换为 `str` 或者 `bytes` 类型的文件系统路径; `os.fsdecode()` 和 `os.fsencode()` 可被分别用来确保获得 `str` 或 `bytes` 类型的结果。此对象是由 [PEP 519](#) 引入的。

**PEP** “Python 增强提议” 的英文缩写。一个 PEP 就是一份设计文档, 用来向 Python 社区提供信息, 或描述一个 Python 的新增特性及其进度或环境。PEP 应当提供精确的技术规格和所提议特性的原理说明。

PEP 应被作为提出主要新特性建议、收集社区对特定问题反馈以及为必须加入 Python 的设计决策编写文档的首选机制。PEP 的作者有责任在社区内部建立共识, 并应将不同意见也记入文档。

参见 [PEP 1](#)。

**portion – 部分** 构成一个命名空间包的单个目录内文件集合 (也可能存放于一个 `zip` 文件内), 具体定义见 [PEP 420](#)。

**positional argument – 位置参数** 参见 *argument*。

**provisional API – 暂定 API** 暂定 API 是指被有意排除在标准库的向后兼容性保证之外的应用编程接口。虽然此类接口通常不会再有重大改变, 但只要其被标记为暂定, 就可能在核心开发者确定有必要的情况下进行向后不兼容的更改 (甚至包括移除该接口)。此种更改并不会随意进行 – 仅在 API 被加入之前未考虑到的严重基础性缺陷被发现时才可能会这样做。

即便是对暂定 API 来说, 向后不兼容的更改也会被视为 “最后的解决方案” —— 任何问题被确认时都会尽可能先尝试找到一种向后兼容的解决方案。

这种处理过程允许标准库持续不断地演进, 不至于被有问题的长期性设计缺陷所困。详情见 [PEP 411](#)。

**provisional package – 暂定包** 参见 *provisional API*。

**Python 3000** Python 3.x 发布路线的昵称 (这个名字在版本 3 的发布还遥遥无期的时候就已出现了)。有时也被缩写为 “Py3k”。

**Pythonic** 指一个思路或一段代码紧密遵循了 Python 语言最常用的风格和理念, 而不是使用其他语言中通用的概念来实现代码。例如, Python 的常用风格是使用 `for` 语句循环来遍历一个可迭代对象中的所有元素。许多其他语言没有这样的结构, 因此不熟悉 Python 的人有时会选择使用一个数字计数器:

```
for i in range(len(food)):
    print(food[i])
```

而相应的更简洁更 Pythonic 的方法是这样的:

```
for piece in food:
    print(piece)
```

**qualified name – 限定名称** 一个以点号分隔的名称，显示从模块的全局作用域到该模块中定义的某个类、函数或方法的“路径”，相关定义见 [PEP 3155](#)。对于最高层级的函数和类，限定名称与对象名称一致：

```
>>> class C:
...     class D:
...         def meth(self):
...             pass
...
>>> C.__qualname__
'C'
>>> C.D.__qualname__
'C.D'
>>> C.D.meth.__qualname__
'C.D.meth'
```

当被用于引用模块时，完整限定名称意为标示该模块的以点号分隔的整个路径，其中包含其所有的父包，例如 `email.mime.text`：

```
>>> import email.mime.text
>>> email.mime.text.__name__
'email.mime.text'
```

**reference count – 引用计数** 对特定对象的引用的数量。当一个对象的引用计数降为零时，所分配资源将被释放。引用计数对 Python 代码来说通常是不可见的，但它是 *CPython* 实现的一个关键元素。`sys` 模块定义了一个 `getrefcount()` 函数，程序员可调用它来返回特定对象的引用计数。

**regular package – 正规包** 传统型的 *package*，例如包含有一个 `__init__.py` 文件的目录。

另参见 *namespace package*。

**\_\_slots\_\_** 一种写在类内部的声明，通过预先声明实例属性等对象并移除实例字典来节省内存。虽然这种技巧很流行，但想要用好却并不容易，最好是只保留在少数情况下采用，例如极耗内存的应用程序，并且其中包含大量实例。

**sequence – 序列** 一种 *iterable*，它支持通过 `__getitem__()` 特殊方法来使用整数索引进行高效的元素访问，并定义了一个返回序列长度的 `__len__()` 方法。内置的序列类型有 `list`、`str`、`tuple` 和 `bytes`。注意虽然 `dict` 也支持 `__getitem__()` 和 `__len__()`，但它被认为属于映射而非序列，因为它查找时使用任意的 *immutable* 键而非整数。

`collections.abc.Sequence` 抽象基类定义了一个更丰富的接口，它超越了 `__getitem__()` 和 `__len__()`，添加了 `count()`、`index()`、`__contains__()` 和 `__reversed__()`。可以使用 `register()` 显式注册实现此扩展接口的类型。

**single dispatch – 单分派** 一种 *generic function* 分派形式，其实现是基于单个参数的类型来选择的。

**slice – 切片** 通常只包含了特定 *sequence* 的一部分的对象。切片是通过使用下标标记来创建的，在 `[]` 中给出几个以冒号分隔的数字，例如 `variable_name[1:3:5]`。方括号（下标）标记在内部使用 *slice* 对象。

**special method – 特殊方法** 一种由 Python 隐式调用的方法，用来对某个类型执行特定操作例如相加等等。这种方法的名称的首尾都为双下划线。特殊方法的文档参见 `specialnames`。

**statement – 语句** 语句是程序段（一个代码“块”）的组成单位。一条语句可以是一个 *expression* 或某个带有关键字的结构，例如 `if`、`while` 或 `for`。

**struct sequence – 结构序列** 具有命名元素的元组。结构序列所暴露的接口类似于 *named tuple*，其元素既可通过索引也可作为属性来访问。不过，它们没有任何具名元组的方法，例如 `_make()` 或 `_asdict()`。结构序列的例子包括 `sys.float_info` 以及 `os.stat()` 的返回值。

**text encoding – 文本编码** 用于将 Unicode 字符串编码为字节的编码器。



**text file – 文本文件** 一种能够读写 `str` 对象的 *file object*。通常一个文本文件实际是访问一个面向字节的数据流并自动处理 *text encoding*。文本文件的例子包括以文本模式 ('r' 或 'w') 打开的文件、`sys.stdin`、`sys.stdout` 以及 `io.StringIO` 的实例。

另请参看 *binary file* 了解能够读写字节类对象的文件对象。

**triple-quoted string – 三引号字符串** 首尾各带三个连续双引号 (") 或者单引号 (') 的字符串。它们在功能上与首尾各用一个引号标注的字符串没有什么不同，但是有多种用处。它们允许你在字符串内包含未经转义的单引号和双引号，并且可以跨越多行而无需使用连接符，在编写文档字符串时特别好用。

**type – 类型** 类型决定一个 Python 对象属于什么种类；每个对象都具有一种类型。要知道对象的类型，可以访问它的 `__class__` 属性，或是通过 `type(obj)` 来获取。

**type alias – 类型别名** 一个类型的同义词，创建方式是把类型赋值给特定的标识符。

类型别名的作用是简化类型提示。例如：

```
from typing import List, Tuple

def remove_gray_shades(
    colors: List[Tuple[int, int, int]]) -> List[Tuple[int, int, int]]:
    pass
```

可以这样提高可读性：

```
from typing import List, Tuple

Color = Tuple[int, int, int]

def remove_gray_shades(colors: List[Color]) -> List[Color]:
    pass
```

参见 `typing` 和 **PEP 484**，其中有对此功能的详细描述。

**type hint – 类型提示** *annotation* 为变量、类属性、函数的形参或返回值指定预期的类型。

类型提示属于可选项，Python 不要求提供，但其可对静态类型分析工具起作用，并可协助 IDE 实现代码补全与重构。

全局变量、类属性和函数的类型提示可以使用 `typing.get_type_hints()` 来访问，但局部变量则不可以。

参见 `typing` 和 **PEP 484**，其中有对此功能的详细描述。

**universal newlines – 通用换行** 一种解读文本流的方式，将以下所有符号都识别为行结束标志：Unix 的行结束约定 '\n'、Windows 的约定 '\r\n' 以及旧版 Macintosh 的约定 '\r'。参见 **PEP 278** 和 **PEP 3116** 和 `bytes.splitlines()` 了解更多用法说明。

**variable annotation – 变量标注** 对变量或类属性的 *annotation*。

在标注变量或类属性时，还可选择为其赋值：

```
class C:
    field: 'annotation'
```

变量标注通常被用作类型提示：例如以下变量预期接受 `int` 类型的值：

```
count: int = 0
```

变量标注语法的详细解释见 *annassign* 一节。

请参看 *function annotation*、**PEP 484** 和 **PEP 526**，其中对此功能有详细描述。

**virtual environment – 虚拟环境** 一种采用协作式隔离的运行时环境，允许 Python 用户和应用程序在安装和升级 Python 分发包时不会干扰到同一系统上运行的其他 Python 应用程序的行为。

另参见 `venv`。

**virtual machine – 虚拟机** 一台完全通过软件定义的计算机。Python 虚拟机可执行字节码编译器所生成的 *bytecode*。

**Zen of Python – Python 之禅** 列出 Python 设计的原则与哲学，有助于理解与使用这种语言。查看其具体内容可在交互模式提示符中输入 `import this`。

---

## 文档说明

---

这些文档生成自 [reStructuredText](#) 原文档，由 [Sphinx](#)（一个专门为 Python 文档写的文档生成器）创建。

本文档和它所用工具链的开发完全是由志愿者完成的，这和 Python 本身一样。如果您想参与进来，请阅读 [reporting-bugs](#) 了解如何参与。我们随时欢迎新的志愿者！

非常感谢：

- Fred L. Drake, Jr., 创造了用于早期 Python 文档的工具链，以及撰写了非常多的文档；
- [Docutils](#) 软件包 项目，创建了 [reStructuredText](#) 文本格式和 [Docutils](#) 软件套件；
- Fredrik Lundh, Sphinx 从他的 [Alternative Python Reference](#) 项目中获得了很多好的想法。

## B.1 Python 文档贡献者

有很多对 Python 语言，Python 标准库和 Python 文档有贡献的人，随 Python 源代码发布的 [Misc/ACKS](#) 文件列出了部分贡献者。

有了 Python 社区的输入和贡献，Python 才有了如此出色的文档 - 谢谢你们！



---

## 历史和许可证

---

### C.1 软件历史

Python 作为 ABC 语言的继承者，于 20 世纪 90 年代初由荷兰的 Guido van Rossum 在荷兰数学和计算机科学研究学会（CWI，见 <https://www.cwi.nl/>）创建。Guido 仍然是 Python 的主要作者，尽管它包含了许多其他人的贡献。

1995 年，Guido 在弗吉尼亚州的国家创新研究公司（CNRI，见 <https://www.cnri.reston.va.us/>）继续他的 Python 工作，在那里他发布了该软件的几个版本。

2000 年 5 月，Guido 和 Python 核心开发团队转移到 BeOpen.com，组建了 BeOpen PythonLabs 团队。同年 10 月，PythonLabs 团队转向 Digital Creations（现为 Zope Corporation；见 <http://www.zope.com/>）。2001 年，Python 软件基金会（PSF，请参阅 <https://www.python.org/psf/>）成立，这是一个专门为拥有与 Python 相关的知识产权而创建的非营利组织。Zope Corporation 是 PSF 的赞助成员。

所有的 Python 版本都是开源的（有关开源的定义参阅 <https://opensource.org/>）。历史上，绝大多数 Python 版本是 GPL 兼容的；下表总结了各个版本情况。

发布版本	源自	年	所有者	GPL 兼容？
0.9.0 至 1.2	n/a	1991-1995	CWI	yes
1.3 至 1.5.2	1.2	1995-1999	CNRI	yes
1.6	1.5.2	2000	CNRI	no
2.0	1.6	2000	BeOpen.com	no
1.6.1	1.6	2001	CNRI	no
2.1	2.0+1.6.1	2001	PSF	no
2.0.1	2.0+1.6.1	2001	PSF	yes
2.1.1	2.1+2.0.1	2001	PSF	yes
2.1.2	2.1.1	2002	PSF	yes
2.1.3	2.1.2	2002	PSF	yes
2.2 及更高版本	2.1.1	2001 至今	PSF	yes

---

**注解：** GPL 兼容并不意味着我们在 GPL 下发布 Python。与 GPL 不同的是，所有 Python 许可都允许您分发修

---

改后的版本，而不需要对所做的更改进行开源。GPL 兼容的许可证使得可以将 Python 与在 GPL 下发布的其他软件组合在一起；其他则不行。

---

感谢许多在 Guido 指导下工作的外部志愿者，使这些发布成为可能。

## C.2 访问 Python 或以其他方式使用 Python 的条款和条件

### C.2.1 Python 3.7.3 的 PSF 许可协议

1. This LICENSE AGREEMENT is between the Python Software Foundation ("PSF"),  
→and  
the Individual or Organization ("Licensee") accessing and otherwise using  
→Python  
3.8.0b1 software in source or binary form and its associated documentation.
2. Subject to the terms and conditions of this License Agreement, PSF hereby  
grants Licensee a nonexclusive, royalty-free, world-wide license to  
→reproduce,  
analyze, test, perform and/or display publicly, prepare derivative works,  
distribute, and otherwise use Python 3.8.0b1 alone or in any derivative  
version, provided, however, that PSF's License Agreement and PSF's notice  
→of  
copyright, i.e., "Copyright © 2001-2019 Python Software Foundation; All  
→Rights  
Reserved" are retained in Python 3.8.0b1 alone or in any derivative version  
prepared by Licensee.
3. In the event Licensee prepares a derivative work that is based on or  
incorporates Python 3.8.0b1 or any part thereof, and wants to make the  
derivative work available to others as provided herein, then Licensee  
→hereby  
agrees to include in any such work a brief summary of the changes made to  
→Python  
3.8.0b1.
4. PSF is making Python 3.8.0b1 available to Licensee on an "AS IS" basis.  
PSF MAKES NO REPRESENTATIONS OR WARRANTIES, EXPRESS OR IMPLIED. BY WAY OF  
EXAMPLE, BUT NOT LIMITATION, PSF MAKES NO AND DISCLAIMS ANY REPRESENTATION  
→OR  
WARRANTY OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OR THAT  
→THE  
USE OF PYTHON 3.8.0b1 WILL NOT INFRINGE ANY THIRD PARTY RIGHTS.
5. PSF SHALL NOT BE LIABLE TO LICENSEE OR ANY OTHER USERS OF PYTHON 3.8.0b1  
FOR ANY INCIDENTAL, SPECIAL, OR CONSEQUENTIAL DAMAGES OR LOSS AS A RESULT  
→OF  
MODIFYING, DISTRIBUTING, OR OTHERWISE USING PYTHON 3.8.0b1, OR ANY  
→DERIVATIVE  
THEREOF, EVEN IF ADVISED OF THE POSSIBILITY THEREOF.

6. This License Agreement will automatically terminate upon a material breach of its terms and conditions.
7. Nothing in this License Agreement shall be deemed to create any relationship of agency, partnership, or joint venture between PSF and Licensee. This License Agreement does not grant permission to use PSF trademarks or trade name in a trademark sense to endorse or promote products or services of Licensee, or any third party.
8. By copying, installing or otherwise using Python 3.8.0b1, Licensee agrees to be bound by the terms and conditions of this License Agreement.

## C.2.2 Python 2.0 的 BeOpen.com 许可协议

### Beopen Python 开源许可协议第 1 版

1. This LICENSE AGREEMENT is between BeOpen.com ("BeOpen"), having an office at 160 Saratoga Avenue, Santa Clara, CA 95051, and the Individual or Organization ("Licensee") accessing and otherwise using this software in source or binary form and its associated documentation ("the Software").
2. Subject to the terms and conditions of this BeOpen Python License Agreement, BeOpen hereby grants Licensee a non-exclusive, royalty-free, world-wide license to reproduce, analyze, test, perform and/or display publicly, prepare derivative works, distribute, and otherwise use the Software alone or in any derivative version, provided, however, that the BeOpen Python License is retained in the Software, alone or in any derivative version prepared by Licensee.
3. BeOpen is making the Software available to Licensee on an "AS IS" basis. BEOPEN MAKES NO REPRESENTATIONS OR WARRANTIES, EXPRESS OR IMPLIED. BY WAY OF EXAMPLE, BUT NOT LIMITATION, BEOPEN MAKES NO AND DISCLAIMS ANY REPRESENTATION OR WARRANTY OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OR THAT THE USE OF THE SOFTWARE WILL NOT INFRINGE ANY THIRD PARTY RIGHTS.
4. BEOPEN SHALL NOT BE LIABLE TO LICENSEE OR ANY OTHER USERS OF THE SOFTWARE FOR ANY INCIDENTAL, SPECIAL, OR CONSEQUENTIAL DAMAGES OR LOSS AS A RESULT OF USING, MODIFYING OR DISTRIBUTING THE SOFTWARE, OR ANY DERIVATIVE THEREOF, EVEN IF ADVISED OF THE POSSIBILITY THEREOF.
5. This License Agreement will automatically terminate upon a material breach of its terms and conditions.
6. This License Agreement shall be governed by and interpreted in all respects by the law of the State of California, excluding conflict of law provisions. Nothing in this License Agreement shall be deemed to create any relationship of agency, partnership, or joint venture between BeOpen and Licensee. This License Agreement does not grant permission to use BeOpen trademarks or trade names in a trademark sense to endorse or promote products or services of Licensee, or any third party. As an exception, the "BeOpen Python" logos available at <http://www.pythonlabs.com/logos.html> may be used according to the permissions

(下页继续)

(续上页)

granted on that web page.

7. By copying, installing or otherwise using the software, Licensee agrees to be bound by the terms and conditions of this License Agreement.

## C.2.3 Python 1.6.1 的 CNRI 许可协议

1. This LICENSE AGREEMENT is between the Corporation for National Research Initiatives, having an office at 1895 Preston White Drive, Reston, VA 20191 ("CNRI"), and the Individual or Organization ("Licensee") accessing and otherwise using Python 1.6.1 software in source or binary form and its associated documentation.
2. Subject to the terms and conditions of this License Agreement, CNRI hereby grants Licensee a nonexclusive, royalty-free, world-wide license to reproduce, analyze, test, perform and/or display publicly, prepare derivative works, distribute, and otherwise use Python 1.6.1 alone or in any derivative version, provided, however, that CNRI's License Agreement and CNRI's notice of copyright, i.e., "Copyright © 1995-2001 Corporation for National Research Initiatives; All Rights Reserved" are retained in Python 1.6.1 alone or in any derivative version prepared by Licensee. Alternately, in lieu of CNRI's License Agreement, Licensee may substitute the following text (omitting the quotes): "Python 1.6.1 is made available subject to the terms and conditions in CNRI's License Agreement. This Agreement together with Python 1.6.1 may be located on the Internet using the following unique, persistent identifier (known as a handle): 1895.22/1013. This Agreement may also be obtained from a proxy server on the Internet using the following URL: <http://hdl.handle.net/1895.22/1013>."
3. In the event Licensee prepares a derivative work that is based on or incorporates Python 1.6.1 or any part thereof, and wants to make the derivative work available to others as provided herein, then Licensee hereby agrees to include in any such work a brief summary of the changes made to Python 1.6.1.
4. CNRI is making Python 1.6.1 available to Licensee on an "AS IS" basis. CNRI MAKES NO REPRESENTATIONS OR WARRANTIES, EXPRESS OR IMPLIED. BY WAY OF EXAMPLE, BUT NOT LIMITATION, CNRI MAKES NO AND DISCLAIMS ANY REPRESENTATION OR WARRANTY OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OR THAT THE USE OF PYTHON 1.6.1 WILL NOT INFRINGE ANY THIRD PARTY RIGHTS.
5. CNRI SHALL NOT BE LIABLE TO LICENSEE OR ANY OTHER USERS OF PYTHON 1.6.1 FOR ANY INCIDENTAL, SPECIAL, OR CONSEQUENTIAL DAMAGES OR LOSS AS A RESULT OF MODIFYING, DISTRIBUTING, OR OTHERWISE USING PYTHON 1.6.1, OR ANY DERIVATIVE THEREOF, EVEN IF ADVISED OF THE POSSIBILITY THEREOF.
6. This License Agreement will automatically terminate upon a material breach of its terms and conditions.
7. This License Agreement shall be governed by the federal intellectual property law of the United States, including without limitation the federal copyright law, and, to the extent such U.S. federal law does not apply, by the law of the Commonwealth of Virginia, excluding Virginia's conflict of law provisions. Notwithstanding the foregoing, with regard to derivative works based on Python 1.6.1 that incorporate non-separable material that was previously distributed under the GNU General Public License (GPL), the law of the Commonwealth of

(下页继续)



(续上页)

Virginia shall govern this License Agreement only as to issues arising under or with respect to Paragraphs 4, 5, and 7 of this License Agreement. Nothing in this License Agreement shall be deemed to create any relationship of agency, partnership, or joint venture between CNRI and Licensee. This License Agreement does not grant permission to use CNRI trademarks or trade name in a trademark sense to endorse or promote products or services of Licensee, or any third party.

8. By clicking on the "ACCEPT" button where indicated, or by copying, installing or otherwise using Python 1.6.1, Licensee agrees to be bound by the terms and conditions of this License Agreement.

## C.2.4 Python 0.9.0 至 1.2 的 CWI 许可协议

Copyright © 1991 - 1995, Stichting Mathematisch Centrum Amsterdam, The Netherlands. All rights reserved.

Permission to use, copy, modify, and distribute this software and its documentation for any purpose and without fee is hereby granted, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation, and that the name of Stichting Mathematisch Centrum or CWI not be used in advertising or publicity pertaining to distribution of the software without specific, written prior permission.

STICHTING MATHEMATISCH CENTRUM DISCLAIMS ALL WARRANTIES WITH REGARD TO THIS SOFTWARE, INCLUDING ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS, IN NO EVENT SHALL STICHTING MATHEMATISCH CENTRUM BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

## C.3 集成软件的许可和认可

This section is an incomplete, but growing list of licenses and acknowledgements for third-party software incorporated in the Python distribution.

### C.3.1 Mersenne Twister

The `_random` module includes code based on a download from <http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/MT2002/emt19937ar.html>. The following are the verbatim comments from the original code:

A C-program for MT19937, with initialization improved 2002/1/26.  
Coded by Takuji Nishimura and Makoto Matsumoto.

Before using, initialize the state by using `init_genrand(seed)`  
or `init_by_array(init_key, key_length)`.

Copyright (C) 1997 - 2002, Makoto Matsumoto and Takuji Nishimura,

(下页继续)

(续上页)

All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. The names of its contributors may not be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Any feedback is very welcome.

<http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html>

email: m-mat @ math.sci.hiroshima-u.ac.jp (remove space)

## C.3.2 套接字

The socket module uses the functions, `getaddrinfo()`, and `getnameinfo()`, which are coded in separate source files from the WIDE Project, <http://www.wide.ad.jp/>.

Copyright (C) 1995, 1996, 1997, and 1998 WIDE Project.

All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. Neither the name of the project nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE PROJECT AND CONTRIBUTORS ``AS IS'' AND

(下页继续)

(续上页)

ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE PROJECT OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

### C.3.3 Asynchronous socket services

The `asyncio` and `asynchat` modules contain the following notice:

Copyright 1996 by Sam Rushing

All Rights Reserved

Permission to use, copy, modify, and distribute this software and its documentation for any purpose and without fee is hereby granted, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation, and that the name of Sam Rushing not be used in advertising or publicity pertaining to distribution of the software without specific, written prior permission.

SAM RUSHING DISCLAIMS ALL WARRANTIES WITH REGARD TO THIS SOFTWARE, INCLUDING ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS, IN NO EVENT SHALL SAM RUSHING BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

### C.3.4 Cookie management

The `http.cookies` module contains the following notice:

Copyright 2000 by Timothy O'Malley <timo@alum.mit.edu>

All Rights Reserved

Permission to use, copy, modify, and distribute this software and its documentation for any purpose and without fee is hereby granted, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation, and that the name of Timothy O'Malley not be used in advertising or publicity pertaining to distribution of the software without specific, written prior permission.

(下页继续)

(续上页)

```
Timothy O'Malley DISCLAIMS ALL WARRANTIES WITH REGARD TO THIS
SOFTWARE, INCLUDING ALL IMPLIED WARRANTIES OF MERCHANTABILITY
AND FITNESS, IN NO EVENT SHALL Timothy O'Malley BE LIABLE FOR
ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES
WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS,
WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS
ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR
PERFORMANCE OF THIS SOFTWARE.
```

### C.3.5 Execution tracing

The trace module contains the following notice:

```
portions copyright 2001, Autonomous Zones Industries, Inc., all rights...
err... reserved and offered to the public under the terms of the
Python 2.2 license.
Author: Zooko O'Whielacronx
http://zooko.com/
mailto:zooko@zooko.com

Copyright 2000, Mojam Media, Inc., all rights reserved.
Author: Skip Montanaro

Copyright 1999, Bioreason, Inc., all rights reserved.
Author: Andrew Dalke

Copyright 1995-1997, Automatrix, Inc., all rights reserved.
Author: Skip Montanaro

Copyright 1991-1995, Stichting Mathematisch Centrum, all rights reserved.

Permission to use, copy, modify, and distribute this Python software and
its associated documentation for any purpose without fee is hereby
granted, provided that the above copyright notice appears in all copies,
and that both that copyright notice and this permission notice appear in
supporting documentation, and that the name of neither Automatrix,
Bioreason or Mojam Media be used in advertising or publicity pertaining to
distribution of the software without specific, written prior permission.
```

### C.3.6 UUencode and UUdecode functions

The uu module contains the following notice:

```
Copyright 1994 by Lance Ellinghouse
Cathedral City, California Republic, United States of America.
    All Rights Reserved
Permission to use, copy, modify, and distribute this software and its
documentation for any purpose and without fee is hereby granted,
provided that the above copyright notice appear in all copies and that
both that copyright notice and this permission notice appear in
supporting documentation, and that the name of Lance Ellinghouse
not be used in advertising or publicity pertaining to distribution
```

(下页继续)

(续上页)

of the software without specific, written prior permission.  
 LANCE ELLINGHOUSE DISCLAIMS ALL WARRANTIES WITH REGARD TO  
 THIS SOFTWARE, INCLUDING ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND  
 FITNESS, IN NO EVENT SHALL LANCE ELLINGHOUSE CENTRUM BE LIABLE  
 FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES  
 WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN  
 ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT  
 OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

Modified by Jack Jansen, CWI, July 1995:

- Use binascii module to do the actual line-by-line conversion  
 between ascii and binary. This results in a 1000-fold speedup. The C  
 version is still 5 times faster, though.
- Arguments more compliant with Python standard

### C.3.7 XML Remote Procedure Calls

The `xmlrpc.client` module contains the following notice:

The XML-RPC client interface is

Copyright (c) 1999-2002 by Secret Labs AB  
 Copyright (c) 1999-2002 by Fredrik Lundh

By obtaining, using, and/or copying this software and/or its  
 associated documentation, you agree that you have read, understood,  
 and will comply with the following terms and conditions:

Permission to use, copy, modify, and distribute this software and  
 its associated documentation for any purpose and without fee is  
 hereby granted, provided that the above copyright notice appears in  
 all copies, and that both that copyright notice and this permission  
 notice appear in supporting documentation, and that the name of  
 Secret Labs AB or the author not be used in advertising or publicity  
 pertaining to distribution of the software without specific, written  
 prior permission.

SECRET LABS AB AND THE AUTHOR DISCLAIMS ALL WARRANTIES WITH REGARD  
 TO THIS SOFTWARE, INCLUDING ALL IMPLIED WARRANTIES OF MERCHANT-  
 ABILITY AND FITNESS. IN NO EVENT SHALL SECRET LABS AB OR THE AUTHOR  
 BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY  
 DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS,  
 WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS  
 ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE  
 OF THIS SOFTWARE.

### C.3.8 test\_epoll

The `test_epoll` module contains the following notice:

Copyright (c) 2001-2006 Twisted Matrix Laboratories.

Permission is hereby granted, free of charge, to any person obtaining

(下页继续)

(续上页)

a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

### C.3.9 Select kqueue

The select module contains the following notice for the kqueue interface:

Copyright (c) 2000 Doug White, 2006 James Knight, 2007 Christian Heimes  
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE AUTHOR AND CONTRIBUTORS ``AS IS'' AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE AUTHOR OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

### C.3.10 SipHash24

The file Python/pyhash.c contains Marek Majkowski's implementation of Dan Bernstein's SipHash24 algorithm. It contains the following note:

<MIT License>  
Copyright (c) 2013 Marek Majkowski <marek@popcount.org>

(下页继续)

(续上页)

```
Permission is hereby granted, free of charge, to any person obtaining a copy
of this software and associated documentation files (the "Software"), to deal
in the Software without restriction, including without limitation the rights
to use, copy, modify, merge, publish, distribute, sublicense, and/or sell
copies of the Software, and to permit persons to whom the Software is
furnished to do so, subject to the following conditions:
```

```
The above copyright notice and this permission notice shall be included in
all copies or substantial portions of the Software.
</MIT License>
```

Original location:

<https://github.com/majek/csiphash/>

Solution inspired by code from:

Samuel Neves (supercop/crypto\_auth/siphash24/little)  
 djb (supercop/crypto\_auth/siphash24/little2)  
 Jean-Philippe Aumasson (<https://131002.net/siphash/siphash24.c>)

### C.3.11 strtod and dtoa

The file `Python/dtoa.c`, which supplies C functions `dtoa` and `strtod` for conversion of C doubles to and from strings, is derived from the file of the same name by David M. Gay, currently available from <http://www.netlib.org/fp/>. The original file, as retrieved on March 16, 2009, contains the following copyright and licensing notice:

```
/*****
 *
 * The author of this software is David M. Gay.
 *
 * Copyright (c) 1991, 2000, 2001 by Lucent Technologies.
 *
 * Permission to use, copy, modify, and distribute this software for any
 * purpose without fee is hereby granted, provided that this entire notice
 * is included in all copies of any software which is or includes a copy
 * or modification of this software and in all copies of the supporting
 * documentation for such software.
 *
 * THIS SOFTWARE IS BEING PROVIDED "AS IS", WITHOUT ANY EXPRESS OR IMPLIED
 * WARRANTY. IN PARTICULAR, NEITHER THE AUTHOR NOR LUCENT MAKES ANY
 * REPRESENTATION OR WARRANTY OF ANY KIND CONCERNING THE MERCHANTABILITY
 * OF THIS SOFTWARE OR ITS FITNESS FOR ANY PARTICULAR PURPOSE.
 *****/
```

### C.3.12 OpenSSL

The modules `hashlib`, `posix`, `ssl`, `crypt` use the OpenSSL library for added performance if made available by the operating system. Additionally, the Windows and Mac OS X installers for Python may include a copy of the OpenSSL libraries, so we include a copy of the OpenSSL license here:

```
LICENSE ISSUES
=====
```

(下页继续)

(续上页)

The OpenSSL toolkit stays under a dual license, i.e. both the conditions of the OpenSSL License and the original SSLeay license apply to the toolkit. See below for the actual license texts. Actually both licenses are BSD-style Open Source licenses. In case of any license issues related to OpenSSL please contact openssl-core@openssl.org.

#### OpenSSL License

-----

```
/* =====
 * Copyright (c) 1998-2008 The OpenSSL Project. All rights reserved.
 *
 * Redistribution and use in source and binary forms, with or without
 * modification, are permitted provided that the following conditions
 * are met:
 *
 * 1. Redistributions of source code must retain the above copyright
 * notice, this list of conditions and the following disclaimer.
 *
 * 2. Redistributions in binary form must reproduce the above copyright
 * notice, this list of conditions and the following disclaimer in
 * the documentation and/or other materials provided with the
 * distribution.
 *
 * 3. All advertising materials mentioning features or use of this
 * software must display the following acknowledgment:
 * "This product includes software developed by the OpenSSL Project
 * for use in the OpenSSL Toolkit. (http://www.openssl.org/)"
 *
 * 4. The names "OpenSSL Toolkit" and "OpenSSL Project" must not be used to
 * endorse or promote products derived from this software without
 * prior written permission. For written permission, please contact
 * openssl-core@openssl.org.
 *
 * 5. Products derived from this software may not be called "OpenSSL"
 * nor may "OpenSSL" appear in their names without prior written
 * permission of the OpenSSL Project.
 *
 * 6. Redistributions of any form whatsoever must retain the following
 * acknowledgment:
 * "This product includes software developed by the OpenSSL Project
 * for use in the OpenSSL Toolkit (http://www.openssl.org/)"
 *
 * THIS SOFTWARE IS PROVIDED BY THE OpenSSL PROJECT ``AS IS'' AND ANY
 * EXPRESSED OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE
 * IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR
 * PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE OpenSSL PROJECT OR
 * ITS CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL,
 * SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT
 * NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES;
 * LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION)
 * HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT,
 * STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE)
 * ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED
 * OF THE POSSIBILITY OF SUCH DAMAGE.
```

(下页继续)



(续上页)

```

* =====
*
* This product includes cryptographic software written by Eric Young
* (eay@cryptsoft.com). This product includes software written by Tim
* Hudson (tjh@cryptsoft.com).
*
*/

```

Original SSLeay License

```

/* Copyright (C) 1995-1998 Eric Young (eay@cryptsoft.com)
* All rights reserved.
*
* This package is an SSL implementation written
* by Eric Young (eay@cryptsoft.com).
* The implementation was written so as to conform with Netscapes SSL.
*
* This library is free for commercial and non-commercial use as long as
* the following conditions are aheared to. The following conditions
* apply to all code found in this distribution, be it the RC4, RSA,
* lhash, DES, etc., code; not just the SSL code. The SSL documentation
* included with this distribution is covered by the same copyright terms
* except that the holder is Tim Hudson (tjh@cryptsoft.com).
*
* Copyright remains Eric Young's, and as such any Copyright notices in
* the code are not to be removed.
* If this package is used in a product, Eric Young should be given attribution
* as the author of the parts of the library used.
* This can be in the form of a textual message at program startup or
* in documentation (online or textual) provided with the package.
*
* Redistribution and use in source and binary forms, with or without
* modification, are permitted provided that the following conditions
* are met:
* 1. Redistributions of source code must retain the copyright
* notice, this list of conditions and the following disclaimer.
* 2. Redistributions in binary form must reproduce the above copyright
* notice, this list of conditions and the following disclaimer in the
* documentation and/or other materials provided with the distribution.
* 3. All advertising materials mentioning features or use of this software
* must display the following acknowledgement:
* "This product includes cryptographic software written by
* Eric Young (eay@cryptsoft.com)"
* The word 'cryptographic' can be left out if the rouines from the library
* being used are not cryptographic related :-).
* 4. If you include any Windows specific code (or a derivative thereof) from
* the apps directory (application code) you must include an acknowledgement:
* "This product includes software written by Tim Hudson (tjh@cryptsoft.com)"
*
* THIS SOFTWARE IS PROVIDED BY ERIC YOUNG ``AS IS'' AND
* ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE
* IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE
* ARE DISCLAIMED. IN NO EVENT SHALL THE AUTHOR OR CONTRIBUTORS BE LIABLE
* FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL
* DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS

```

(下页继续)

(续上页)

```
* OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION)
* HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT
* LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY
* OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF
* SUCH DAMAGE.
*
* The licence and distribution terms for any publically available version or
* derivative of this code cannot be changed. i.e. this code cannot simply be
* copied and put under another distribution licence
* [including the GNU Public Licence.]
*/
```

### C.3.13 expat

The pyexpat extension is built using an included copy of the expat sources unless the build is configured `--with-system-expat`:

```
Copyright (c) 1998, 1999, 2000 Thai Open Source Software Center Ltd
and Clark Cooper

Permission is hereby granted, free of charge, to any person obtaining
a copy of this software and associated documentation files (the
"Software"), to deal in the Software without restriction, including
without limitation the rights to use, copy, modify, merge, publish,
distribute, sublicense, and/or sell copies of the Software, and to
permit persons to whom the Software is furnished to do so, subject to
the following conditions:

The above copyright notice and this permission notice shall be included
in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND,
EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF
MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT.
IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY
CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT,
TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE
SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.
```

### C.3.14 libffi

The `_ctypes` extension is built using an included copy of the libffi sources unless the build is configured `--with-system-libffi`:

```
Copyright (c) 1996-2008 Red Hat, Inc and others.

Permission is hereby granted, free of charge, to any person obtaining
a copy of this software and associated documentation files (the
``Software''), to deal in the Software without restriction, including
without limitation the rights to use, copy, modify, merge, publish,
distribute, sublicense, and/or sell copies of the Software, and to
permit persons to whom the Software is furnished to do so, subject to
the following conditions:
```

(下页继续)

(续上页)

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED ``AS IS'', WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

### C.3.15 zlib

The `zlib` extension is built using an included copy of the `zlib` sources if the `zlib` version found on the system is too old to be used for the build:

Copyright (C) 1995-2011 Jean-loup Gailly and Mark Adler

This software is provided 'as-is', without any express or implied warranty. In no event will the authors be held liable for any damages arising from the use of this software.

Permission is granted to anyone to use this software for any purpose, including commercial applications, and to alter it and redistribute it freely, subject to the following restrictions:

1. The origin of this software must not be misrepresented; you must not claim that you wrote the original software. If you use this software in a product, an acknowledgment in the product documentation would be appreciated but is not required.
2. Altered source versions must be plainly marked as such, and must not be misrepresented as being the original software.
3. This notice may not be removed or altered from any source distribution.

Jean-loup Gailly  
jloup@gzip.org

Mark Adler  
madler@alumni.caltech.edu

### C.3.16 cfuhash

The implementation of the hash table used by the `tracemalloc` is based on the `cfuhash` project:

Copyright (c) 2005 Don Owens  
All rights reserved.

This code is released under the BSD license:

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

(下页继续)

(续上页)

- \* Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- \* Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- \* Neither the name of the author nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

### C.3.17 libmpdec

The `_decimal` module is built using an included copy of the libmpdec library unless the build is configured `--with-system-libmpdec`:

Copyright (c) 2008–2016 Stefan Krah. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE AUTHOR AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE AUTHOR OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

### C.3.18 W3C C14N test suite

The C14N 2.0 test suite in the `test` package (`Lib/test/xmltestdata/c14n-20/`) was retrieved from the W3C website at <https://www.w3.org/TR/xml-c14n2-testcases/> and is distributed under the 3-clause BSD license:

Copyright (c) 2013 W3C(R) (MIT, ERCIM, Keio, Beihang), All Rights Reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of works must retain the original copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the original copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- Neither the name of the W3C nor the names of its contributors may be used to endorse or promote products derived from this work without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



---

### 版权

---

Python 和这份文档的版权说明：

Copyright © 2001-2019 Python Software Foundation. All rights reserved.

Copyright © 2000 BeOpen.com. All rights reserved.

Copyright © 1995-2000 Corporation for National Research Initiatives. All rights reserved.

Copyright © 1991-1995 Stichting Mathematisch Centrum. All rights reserved.

---

更完整的版权信息，请参考[历史和许可证](#)。





## 非字母

..., [67](#)  
 2to3, [67](#)  
 >>>, [67](#)  
 \_\_future\_\_, [71](#)  
 \_\_slots\_\_, [76](#)  
 环境变量  
     PYTHONPATH, [55](#)

## A

abstract base class -- 抽象基类, [67](#)  
 annotation -- 标注, [67](#)  
 argument -- 参数, [67](#)  
 asynchronous context manager -- 异步上下  
     文管理器, [68](#)  
 asynchronous generator -- 异步生成器, [68](#)  
 asynchronous generator iterator -- 异  
     步生成器迭代器, [68](#)  
 asynchronous iterable -- 异步可迭代对象,  
     [68](#)  
 asynchronous iterator -- 异步迭代器, [68](#)  
 attribute -- 属性, [68](#)  
 awaitable -- 可等待对象, [68](#)

## B

BDFL, [68](#)  
 binary file -- 二进制文件, [68](#)  
 bytecode -- 字节码, [68](#)  
 bytes-like object -- 字节类对象, [68](#)

## C

C-contiguous, [69](#)  
 class -- 类, [69](#)  
 class variable -- 类变量, [69](#)  
 coercion -- 强制类型转换, [69](#)  
 complex number -- 复数, [69](#)  
 context manager -- 上下文管理器, [69](#)  
 context variable -- 上下文变量, [69](#)  
 contiguous -- 连续, [69](#)

coroutine -- 协程, [69](#)  
 coroutine function -- 协程函数, [69](#)  
 CPython, [69](#)

## D

deallocation, object, [47](#)  
 decorator -- 装饰器, [69](#)  
 descriptor -- 描述器, [69](#)  
 dictionary -- 字典, [70](#)  
 dictionary view -- 字典视图, [70](#)  
 docstring -- 文档字符串, [70](#)  
 duck-typing -- 鸭子类型, [70](#)

## E

EAFP, [70](#)  
 expression -- 表达式, [70](#)  
 extension module -- 扩展模块, [70](#)

## F

file object -- 文件对象, [70](#)  
 file-like object -- 文件类对象, [70](#)  
 finalization, of objects, [47](#)  
 finder -- 查找器, [70](#)  
 floor division -- 向下取整除法, [70](#)  
 Fortran contiguous, [69](#)  
 f-string -- f-字符串, [70](#)  
 function -- 函数, [70](#)  
 function annotation -- 函数标注, [70](#)

## G

garbage collection -- 垃圾回收, [71](#)  
 generator, [71](#)  
 generator -- 生成器, [71](#)  
 generator expression, [71](#)  
 generator expression -- 生成器表达式, [71](#)  
 generator iterator -- 生成器迭代器, [71](#)  
 generic function -- 泛型函数, [71](#)  
 GIL, [71](#)

global interpreter lock -- 全局解释器锁, 71

## H

hashable -- 可哈希, 71  
hash-based pyc -- 基于哈希的 pyc, 71

## I

IDLE, 72  
immutable -- 不可变, 72  
import path -- 导入路径, 72  
importer -- 导入器, 72  
importing -- 导入, 72  
interactive -- 交互, 72  
interpreted -- 解释型, 72  
interpreter shutdown -- 解释器关闭, 72  
iterable -- 可迭代对象, 72  
iterator -- 迭代器, 72

## K

key function -- 键函数, 72  
keyword argument -- 关键字参数, 73

## L

lambda, 73  
LBYL, 73  
list -- 列表, 73  
list comprehension -- 列表推导式, 73  
loader -- 加载器, 73

## M

magic  
    method, 73  
magic method -- 魔术方法, 73  
mapping -- 映射, 73  
meta path finder -- 元路径查找器, 73  
metaclass -- 元类, 73  
method  
    magic, 73  
    special, 76  
method -- 方法, 73  
method resolution order -- 方法解析顺序, 73  
module -- 模块, 73  
module spec -- 模块规格, 73  
MRO, 73  
mutable -- 可变, 74

## N

named tuple -- 具名元组, 74  
namespace -- 命名空间, 74  
namespace package -- 命名空间包, 74  
nested scope -- 嵌套作用域, 74

## O

object  
    deallocation, 47  
    finalization, 47  
object -- 对象, 74

## P

package -- 包, 74  
parameter -- 形参, 74  
path based finder -- 基于路径的查找器, 75  
path entry -- 路径入口, 75  
path entry finder -- 路径入口查找器, 75  
path entry hook -- 路径入口钩子, 75  
path-like object -- 路径类对象, 75  
PEP, 75  
Philbrick, Geoff, 14  
portion -- 部分, 75  
positional argument -- 位置参数, 75  
provisional API -- 暂定 API, 75  
provisional package -- 暂定包, 75  
PyArg\_ParseTuple(), 13  
PyArg\_ParseTupleAndKeywords(), 14  
PyErr\_Fetch(), 48  
PyErr\_Restore(), 48  
PyInit\_modulename (C 函数), 55  
PyObject\_CallObject(), 11  
Python 3000, 75  
Python 提高建议  
    PEP 1, 75  
    PEP 238, 70  
    PEP 278, 77  
    PEP 302, 70, 73  
    PEP 343, 69  
    PEP 362, 68, 75  
    PEP 411, 75  
    PEP 420, 70, 74, 75  
    PEP 442, 48  
    PEP 443, 71  
    PEP 451, 70  
    PEP 484, 67, 71, 77  
    PEP 489, 10, 56  
    PEP 492, 68, 69  
    PEP 498, 70  
    PEP 519, 75  
    PEP 525, 68  
    PEP 526, 67, 77  
    PEP 3116, 77  
    PEP 3155, 76  
Pythonic, 75  
PYTHONPATH, 55

## Q

qualified name -- 限定名称, [76](#)

## R

READ\_RESTRICTED, [51](#)

READONLY, [51](#)

reference count -- 引用计数, [76](#)

regular package -- 正规包, [76](#)

repr

    ☐置函数, [48](#)

RESTRICTED, [51](#)

## S

sequence -- 序列, [76](#)

single dispatch -- 单分派, [76](#)

slice -- 切片, [76](#)

special

    method, [76](#)

special method -- 特殊方法, [76](#)

statement -- 语句, [76](#)

string

    object representation, [48](#)

struct sequence -- 结构序列, [76](#)

## T

text encoding -- 文本编码, [76](#)

text file -- 文本文件, [77](#)

triple-quoted string -- 三引号字符串, [77](#)

type -- 类型, [77](#)

type alias -- 类型别名, [77](#)

type hint -- 类型提示, [77](#)

## U

universal newlines -- 通用换行, [77](#)

## V

variable annotation -- 变量标注, [77](#)

☐置函数

    repr, [48](#)

virtual environment -- 虚拟环境, [78](#)

virtual machine -- 虚拟机, [78](#)

## W

WRITE\_RESTRICTED, [51](#)

## Z

Zen of Python -- Python 之禅, [78](#)