

玩转PYTHON

石秀川 2017-04-05

Awesome Python

Life is short, you need Python.

OUTLINE

- 语言概念
- 基本语法 & 特性
- 进阶——实用特性、语法糖、内置函数、库
- 工具链
- 生态系统
- Gotcha

语言概念

- 解释型语言（脚本语言）
- 基于C的Python解释器
- 动态类型
- 强类型
- 面向过程 + 面向对象 + 函数式

基本语法 & 特性

- 没有分号！没有大括号！
- 变量
- 类型
 - 整形 长整形 复数 双精度浮点数
 - 字符串（str, u, r），字符串不变性——性能问题
 - 丰富的数据结构和操作（list, tuple, dict, set...）
- 三大控制语句（if/else，while/for，~~switch/case~~）
- 函数（+元组 = 多返回值）
- 面向对象
- 异常
- 多继承，C3 BFS
- module, sys.path

```
hello.py  hello_python.py  hello >
2 # -*- coding: utf-8 -*-
3
4 import hello
5
6
7 class HelloSpeaker(object):
8
9     def hello(self):
10         print 'Hello Python!'
11         print type('Hello Python!')
12         raise Exception('oops...')
13
14
15 if __name__ == '__main__':
16     a_list = [1, 2, 3]
17     a_list.append(4)
18     print a_list[:2]
19     a_tuple = (1, 2, 3)
20     a_dict = {
21         'a': 'A',
22         'b': 'B',
23     }
24     for k, v in a_dict.items():
25         print k, v
26
27     speaker = HelloSpeaker()
28     try:
29         speaker.hello()
30     except Exception as e:
31         pass
32
33     for i in range(5):
34         hello.say_hello()
35
36 # vim: set expandtab ts=4 sts=4 sw=4 :
NORMAL  hello_python.py
3
4
5 def say_hello():
6     print u'Hello Python!'
7     print type(u'Hello Python!')
8
9 # vim: set expandtab ts=4 sts=4 sw=4 :
hello.py
```

进阶——实用特性、语法糖、内置函数、库 *pythonic*

hello_python.py

- X if C else Y —— 三元运算
- List Comprehension —— 列表生成
- yield —— 迭代器
- map & reduce & filter
- lambda
- decorator —— 函数装饰器
- function: first class
- String format
 - “%” vs “.format()”
- with关键字 —— context manager
- pdb —— pdb with GUI
- logging —— full functional!
 - level、format、namespace定义
 - rotation（大小，时间等维度）
- f(*args, **kwargs)——不定长参数
- 对象存储机制与自省
- 扩展Python, Python binding

工具链

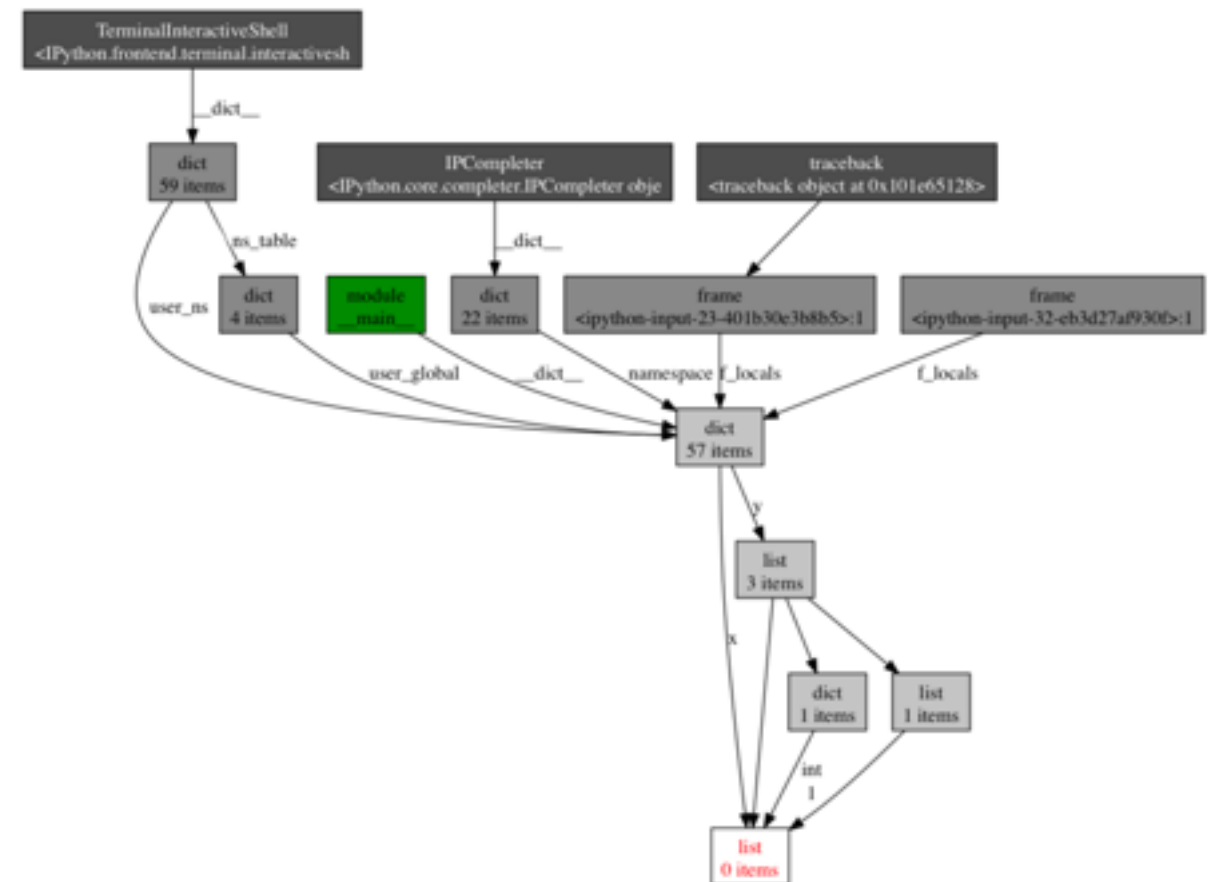
- 编辑器
 - VIM/Emacs
 - PyCharm
- 静态分析 & 代码检查
 - pyflakes
 - pylint
 - PEP 8
- 文档生成与查看
 - docstring
 - pydoc
 - sphinx

工具链

➤ 性能分析

➤ profile/cProfile

➤ line_profiler



Mon Jun 13 20:54:03 2016 timeit

14000087 function calls in 6.604 seconds

Ordered by: internal time

List reduced from 27 to 6 due to restriction <6>

ncalls	tottime	percall	cumtime	percall	filename:lineno(function)
100000	3.234	0.000	4.609	0.000	time_profile.py:11(slowest_replace)
100000	1.693	0.000	1.693	0.000	time_profile.py:20(slow_replace)
13200000	1.180	0.000	1.180	0.000	{method 'append' of 'list' objects}
200000	0.236	0.000	0.236	0.000	{method 'join' of 'str' objects}
100000	0.095	0.000	0.095	0.000	{method 'split' of 'str' objects}
100000	0.072	0.000	0.208	0.000	time_profile.py:28(fast_replace)

工具链

➤ 包管理

- setuptools (setup.py)

- pip, PyPI

- egg vs wheel

- Python 包管理工具解惑

- ZingFront Python common lib: zf_py_lib

工具链

- Virtualenv
 - A tool to create isolated Python environments
 - 解决的问题：
 - 公用环境中，对同一lib的不同版本的依赖
 - 新版本/新工具的实验，新实践探索

生态系统

- Web/Net
 - Django
 - requests等一众net lib
- (Data)Science
 - numpy
 - 一众ML/DL框架 (Caffe, TensorFlow, Keras, etc.)
 - 一众CV/MV的Python binding (OpenCV等)
 - Visualization (Matplotlib等)
- Web Crawler
 - Scrapy
- Message Queue
 - Celery
- Service
 - Thrift
 - gRPC
- Template Engine
 - Jinja2
- DevOps
 - Supervisor
- Processing
 - Imagery (pillow)
 - Office (xlwt + xlrd)

GOTCHA

- 文件编码: ASCII & UTF-8
- “字符串”: unicode & str
- “多”线程:
 - Python等解释型语言, 是“伪多线程”
 - 原因: GIL, 全局解释锁
 - I/O密集✓ vs CPU密集✗
 - 多线程 vs 多进程 (multiprocessing包)

Q & A

.....