《算法与程序实现的综合运用一》拓展资源

1. Python运行环境

电脑：

如果没有第三方库thonny.org，要下载安装thonny。

官网www.python.org 或者 python123.io 安装32位python3.5。

海龟编辑器：<https://python.codemao.cn/>，在“库管理”中选择安装好jieba等库，就可以运行了，例如：

import jieba

y=jieba.lcut("中华民族是一个伟大勤劳勇敢善良的民族")

print(y)

运行后可正确输出分词结果，同学们可以试一试用它做练习。

手机：

python3IDE

网页编程：

每颗豆[www.makerbean.com](http://www.makerbean.com)或者 python123.io

1. 学习资源

菜鸟<https://www.runoob.com>，找python3

微信小程序：懒人学python

强烈推荐：python123.io 或搜“嵩天”，看python视频

1. 爬取网页文本

#爬取小说《西游记》rzq

import requests #requests库可以获取网页内容

from bs4 import BeautifulSoup #bs4库解析网页结构

def getHTMLtext(url): #获取网页url的内容

try:

r = requests.get(url,timeout=30)

r.raise\_for\_status()

r.encoding = r.apparent\_encoding

return r.text

except:

return "出错" #网站访问出错

with open('西游记.txt','w',encoding='utf-8') as f: #建立文件

f.write('西游记\n\n') #写入标题和两个空行

for k in range(100):

#《西游记》有100回，观察到网站上的每回是一个页面

url='http://www.shicimingju.com/book/xiyouji/'+str(k+1)+'.html'

soup=BeautifulSoup(getHTMLtext(url)) #解析一回的结构

f.write('\n')

f.write(soup.find('h1').string)

#观察到每一回的题目在<h1></h1>中

f.write('\n\n')

for i in soup.find\_all('p'): #观察到每一段落在<p></p>中

f.write(i.string)

f.write('\n')