import urllib2

response=urllib2.urlopen("http://www.baidu.com")

html=response.resd()

print html

-------------------------------------------------------------------------

import urllib2

req=urllib2.request("http://www.baidu.com")

response=urllib2.urlopen("")

html=response.resd()

print html

第五周 学习用pythono制作爬虫

存入python2.exe的路径

Rul即为网页的地址

爬虫就是把URL地址中的网络资源读取出来，然后处理

这里使用的是：python2.7 版本

准备：安装python2.7→右键“此电脑”属性→高级设置→环境变量→编辑path环境变量→

存入一个代码为以下四行的文件 命名为pc\_1.py

import urllib2

#此行代码是引入urllib2的库

r= urllib2.urlopen('http://www.baidu.com')

#打开百度网址 并将其命名为一个变量

h= r.read()

#解读该变量存入的读取的网址

print h

#打印该网址用爬虫爬取的结果

打开电脑 win+r 运行 cmd 打入python2.exe 以黑色文本框呼出python2.7即可快捷使用

此时 呼出cmd 输入 pc\_1.py则运行此程序

得到如下:

<!DOCTYPE html><!--STATUS OK--><html><head><meta http-equiv="conten

t-type" content="text/html;charset=utf-8"><meta http-equiv="X-UA-Co

mpatible" content="IE=Edge"><meta content="always" name="referrer">

<meta name="theme-color" content="#2932e1"><link rel="shortcut ico

n" href="/favicon.ico" type="image/x-icon" /><link rel="search" typ

e="application/opensearchdescription+xml" href="/content-search.xm

l" title="百度搜索" /><link rel="icon" sizes="any" mask href="//www.b

aidu.com/img/baidu.svg"><link rel="dns-prefetch" href="//s1.bdstati

c.com"/><link rel="dns-prefetch" href="//t1.baidu.com"/><link re

l="dns-prefetch" href="//t2.baidu.com"/><link rel="dns-prefetch" hr

ef="//t3.baidu.com"/><link rel="dns-prefetch" href="//t10.baidu.co

m"/><link rel="dns-prefetch" href="//t11.baidu.com"/><link rel="dn

s-prefetch" href="//t12.baidu.com"/><link rel="dns-prefetch" hre

f="//b1.bdstatic.com"/><title>百度一下，你就知道</title>

该文本效果如同在浏览器的该网址界面按f12查看的源码相同