Python+Selenium笔记（一）：环境配置+简单的例子

#环境配置基于windows操作系统

#学习selenium要有一些HTML和xpth的基础，完全不会的建议先花点时间学点基础（不然元素定位，特别是xpth可能看的有点懵）

#HTML :  http://www.runoob.com/html/     #xpth:  http://www.runoob.com/xpath

（一）  环境配置

（1）执行下面的命令（前提：已经安装python环境，可以参考之前发的python笔记（一））

pip install -U selenium

（2）http://docs.seleniumhq.org/download/   （网站打不开的话就FQ）

去下载Mozilla GeckoDriver（因为我暂时用火狐测试，所以下载这个），解压后放到任意目录下，然后在系统变量path中加上解压后的路径（我在2台电脑上试了下，一个放在任意目录，一个放到python的安装目录，反正都能用。）

（3）之后如果创建浏览器驱动实例还是报错的话（可能是浏览器版本的问题，我python3.6+selenium3.9+火狐47+Mozilla GeckoDriverV0.19.1就报错，火狐换成58就不会了）

（二）  简单的例子

#导入webdriver  
from selenium import webdriver  
#FirefoxBinary--导入2进制文件所在的位置  
from selenium.webdriver.firefox.firefox\_binary import FirefoxBinary  
#指定火狐浏览器的二进制路径  
firefox = FirefoxBinary(r"C:\Program Files (x86)\Mozilla  
Firefox\firefox.exe")  
driver = webdriver.Firefox(firefox\_binary=firefox) #创建火狐浏览器的驱动实例  
  
driver.implicitly\_wait(2)  
#设置等待时间（定义执行步骤的超时时间）  
driver.maximize\_window()  #最大化浏览器

driver.get("https://www.cnblogs.com/") #打开博客园首页  
search\_field =  
driver.find\_element\_by\_id('zzk\_q') #通过id定位博客园首页的搜索框  
search\_btn  
=driver.find\_element\_by\_class\_name('search\_btn') #通过class定位博客园首页的查询按钮  
  
search\_field.clear() #清空搜索框的值（如果有）  
search\_field.send\_keys("python") #在搜索框中输入查询条件  
search\_btn.click() #点击查询按钮  
# search\_field.submit()  
#选取<h3><a>开始 ，</a>结束之间的值，返回一个列表  
products =  
driver.find\_elements\_by\_xpath('//h3[a/@target="\_blank"]/a')  
#迭代列表products，输出text的值  
for p in products:  
    print(p.text)  
#选取<a target="\_blank">节点的值，返回一个列表  
pro =  
driver.find\_elements\_by\_xpath('//h3/a[@target="\_blank"]')  
#迭代输出href属性的值  
for p in pro:  
    print(p.get\_attribute('href'))  
#关闭浏览器  
driver.quit()

（三）  输出如下

第二步骤中的xpth其实主要定位的就是下面这些



