# 学**习心得**

--------学号：2018212332 姓名：王持禹丞

代理神马的，果真对于我导论都挂的人一点都不友好！！！

这次也仅完成了单线程跑（真的慢），多线程还在学习中，目前实现的功能有自定义后缀名查询，在域名上加了xx.xx.后缀名的形式。

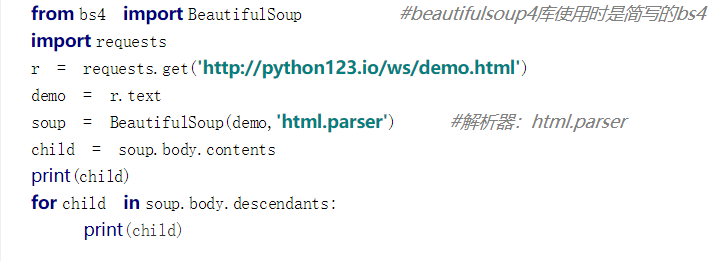
这次脚本我想到的就是爬虫，利用requests爬取注册信息，并将之保存在文件中。

**第一个困难**：在看到一篇博文后，它通过读取13万单词的txt来进行

因此，我必须先写一个能产生4位字母所有组合的的程序，并将其保存在txt文件中。当时想起用random函数，但是跑了一会就发现不能产生所有情况，才意识到直接4个for就可以直接解决。

**第二个困难**：利用requests必须要一个能查.gs的URL,找了好久终于找到万网的，不知道为什么gs都是不可注册，而不是域名已被注册。（后来学长发的那个万网的url，不管怎样就是进不去，自己也去找了许多网站，发现可以查但是就是找不到注册和未注册的状态码）

**第三个困难**：学习bs4，了解了HTML格式，通过



此例子，完成了查询的函数。

**思路想法**：先用requests和beautifulsoup将html中注册和未注册的状态码找到，知道后就可定义domain函数，然后导入4word.py产生的4word.txt，定义damainlist函数，最后将注册与未注册的状态码作为条件成sucessfullist.txt。

这次作业只能说完成基本要求的80%，过程有着很多困惑，但大量尝试以及谷歌之后也懂得七七八八，过程很艰苦，但是达成一个小目标后很有成就感。