```python
#201019_获取全书网正文
#上节课思考题 获取全书网正文
import re
import requests
import os
web_url='https://www.xs4.cc/0_4/'  #选择要爬取的书的网址
web_url2=re.findall('cc/(.*?)/',web_url)[0]  #取得网址的一部分
req=requests.get(web_url)  #实例化一个requests请求
req.encoding='gbk'  #编码设置为gbk,防止乱码
shuming=re.findall('<h1>(.*?)</h1>',req.text)[0]  #取得书名
mulu=re.findall('.html">(.*?)</a></dd>',req.text,re.S)  #获取目录
wangzhi=re.findall(f'<a href="/{web_url2}/(.*?).html"',req.text)  #获取网址的一部分
dict1={}
for i in range(9,len(mulu)):
    dict1[mulu[i]]=f'{web_url}{wangzhi[i]}.html' #把目录和网址存入字典中
if os.path.exists(f'd:/{shuming}'):  #如果目录存在,则不做操作
    pass
else:
    os.mkdir(f'd:/{shuming}')  #如果目录不存在,则创建以书名命名的目录
count=0
for k,v in dict1.items():
    if count>=3:  #控制循环的次数
        break
    else:
        zhengwen=requests.get(v)
        zhengwen.encoding='gbk'
        neirong=re.findall('<div id="content">(.*?)</div>',zhengwen.text,re.S)[0]
        neirong=neirong.replace(' ','').replace('<br />','')
        with open(f'd:/{shuming}/{k}.txt','w+') as file1:
            file1.write(neirong)
        count += 1
```