

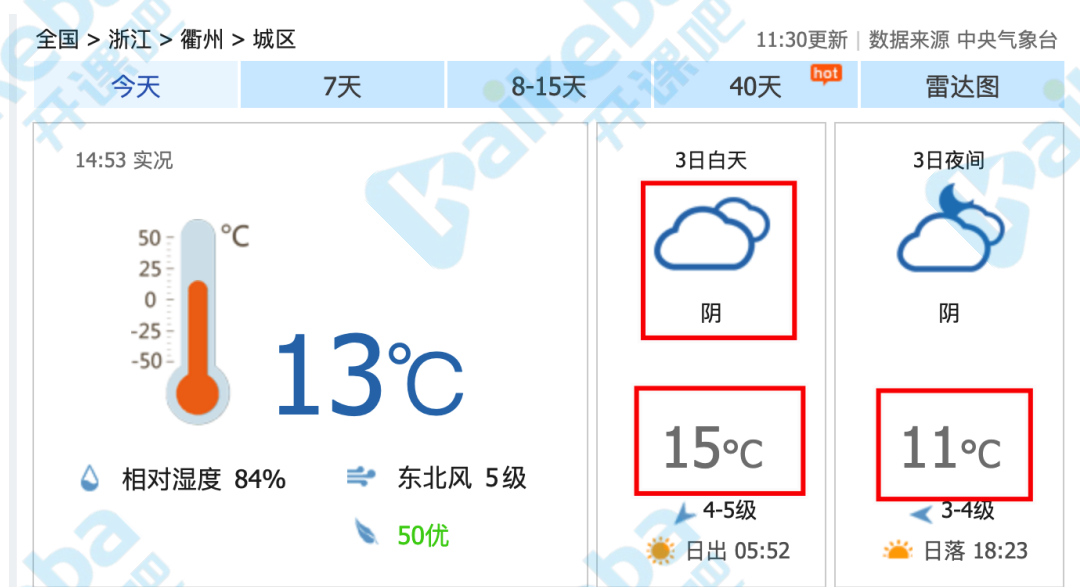
第11课 让爬虫定时向你汇报答案

题目要求:

- 根据已给出的网址爬取衢州的天气, 爬取白天、夜间的温度、阴天或晴天数据具体数据位置见下面图片, 红色框圈出来的是需要爬取的数据。将爬取到的数据写入到发送邮件模板中并定时每天早上八点, 发送天气情况。

使用技术点引导

- requests
- BeautifulSoup
- smtplib
- email
- schedule
- time
- while循环
- 利用函数封装代码 (选用)



```
1 # 导入模块
2 import requests
3 from bs4 import BeautifulSoup
4 import smtplib
5 from email.mime.text import MIMEText
6 from email.header import Header
7 import schedule
8 import time
9
10 # 输入邮箱发件人、收件人以及邮箱的授权码
12 account = str(input('请输入发件人邮箱地址: '))
13 password = str(input('请输入邮箱授权码: '))
```

```
14 receiver = str(input('请输入收件人邮箱地址: '))
15 # 建立天气网爬虫, 爬取天气信息
20 def weather_spider():
21     # 模拟浏览器:
24     headers = {
25         'user-agent' : 'Mozilla/5.0 (Macintosh; Intel Mac OS X 10_13_6)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/71.0.3578.98 Safari/537.36'
26     }
27     url = 'http://www.weather.com.cn/weather/101211001.shtml'
28     # 数据获取:
29     res = requests.get(url, headers=headers)
30     res.encoding = 'utf-8'
31     # 数据解析:
32     soup = BeautifulSoup(res.text, 'html.parser')
33     # 数据提取:
34     tem1 = soup.find(class_='tem')
35     weather1 = soup.find(class_='wea')
36     tem = tem1.text
37     weather = weather1.text
38     print(tem, weather)
39     return tem, weather
40 # weather_spider()
41 # #发送邮件的代码
42 def send_email(tem, weather):
43     mailhost = 'smtp.qq.com'
44     qqmail = smtplib.SMTP()
45     qqmail.connect(mailhost, 25)
46     qqmail.login(account, password)
47     content = '衢州的天气是: \n'+tem+weather
48     message = MIMEText(content, 'plain', 'utf-8')
49     subject = '今日天气预报from python'
50     message['Subject'] = Header(subject, 'utf-8')
51     qqmail.sendmail(account, receiver, message.as_string())
52     print('邮件发送成功')
53     qqmail.quit()
54 #建立任务
55 def job():
56     print('开始一次任务')
57     tem, weather = weather_spider()
58     # print("发送邮件")
59     send_email(tem, weather)
60     print('任务完成')
61 #定时发送
62 schedule.every().friday.at("14:43").do(job)
63 while True:
64     schedule.run_pending()
65     time.sleep(1)
```