

1 代码结构

1.1 环境配置

```
1 import csv
2 from lxml import etree
3 import asyncio, aiohttp # 协程模块
4 import psychopg
5
6 asyncio.set_event_loop_policy(asyncio.WindowsSelectorEventLoopPolicy()) # windowsSQL协程设置
7
8 headers = {
9     'Referer': 'https://music.163.com/',
10    'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/68.0.3440.106 Safari/537.36'
11 }
```

1.2 协程函数

```
14 async def request(url, headers = headers):
15     conn=aiohttp.TCPConnector(ssl=False)
16     async with aiohttp.ClientSession(connector=conn) as session:
17         response = await session.get(url, headers = headers)
18         return await response.text()
19
20 async def a_insert(row):
21     async with await psychopg.AsyncConnection.connect('dbname=week15 user=postgres password=000000') as conn:
22         async with conn.cursor() as cur:
23             await cur.execute('INSERT INTO playLists VALUES (%s, %s, %s, %s, %s, %s, %s)', row)
24
25
26 async def a_write(response, i):
27     html = etree.HTML(response)
28     ID = html.xpath('//a[@data-res-type="13"]/@data-res-id')[0]
29     title = html.xpath('//h2[@class="f-ff2 f-brk"]')[0].text
30     intro_result = html.xpath('//p[@class="intr f-brk"]')
31     if intro_result == []:
32         text = ''
33     else:
34         text = intro_result[0].xpath('string(.)')[1:-1].replace('\n', ' ') # 歌单标签
35     author = html.xpath('//a[@class="s-fc7"]')[0].text
36     date = html.xpath('//span[@class="time s-fc4"]')[0].text[:10]
37     mp3_path = 'https://music.163.com' + html.xpath('//a[@class="des s-fc7"]/@data-href')[0] # 歌曲外链以备下载
38     related_articles = html.xpath('//p[@class="intr f-brk"]')[0].xpath('string(.)')[4:] # 歌单介绍
39     row = (ID, title, text, author, date, mp3_path, related_articles)
40     await a_insert(row)
41     print(f'>>>--第{i+1}个歌单写入成功--v')
```

1.2 主函数

```
44 if __name__ == '__main__':
45     with open(r'week15_data.csv', 'a', encoding = 'utf-8', newline = '') as f:
46         csv.writer(f).writerow(['ID', 'title', 'text', 'author', 'date', 'mp3_path', 'related_articles'])
47     with open('week15_pl_urls.csv', 'r', encoding = 'utf-8', newline = '') as f:
48         urls = [i[0] for i in csv.reader(f)]
49     loop = asyncio.get_event_loop()
50     tasks = []
51     for i in range(len(urls)):
52         print(f'---开始爬取第{i+1}个歌单-->>>')
53         url = urls[i]
54         task_request = loop.create_task(request(url))
55         loop.run_until_complete(task_request)
56         tasks.append(loop.create_task(a_write(task_request.result(), i)))
57     loop.run_until_complete(asyncio.wait(tasks))
```

2 运行结果

```
E:\北航\课业\大三上\现代程序设计技术\week15>python week15_homework.py
--开始爬取第1个歌单-->>>
--开始爬取第2个歌单-->>>
>>>--第1个歌单写入成功--✓
--开始爬取第3个歌单-->>>
>>>--第2个歌单写入成功--✓
--开始爬取第4个歌单-->>>
>>>--第3个歌单写入成功--✓
--开始爬取第5个歌单-->>>
>>>--第4个歌单写入成功--✓
```

可以看到，爬取和写入文件不同步进行。

写入数据库如下：

	ID [PK] character	title character varying (100)	text character varying (30)	author character varying (20)	date character (10)	mp3_path character varying (200)	related_articles character varying (200)
1	100295664	中文说唱 我要的是一种态度	标签： 说唱 华语 放松	夢遊樹	2015-08-22	https://music.163.co...	喜欢有自己态度的Hiphop 不要听流水账
2	102445100	让你停不下来的带感HipHop	标签： 欧美 说唱 流行	JediMindTricks	2015-08-29	https://music.163.co...	个人平时所听的各种曲风hiphop，新老
3	104533395	「欧美说唱」殿堂级欧美hip hop	标签： 欧美 说唱 蓝调	Leonsamiz	2015-09-06	https://music.163.co...	说唱(RAP)， 20世纪90年代初风靡欧美
4	107203462	Gfunk & Funk & Chicano Rap ...	标签： 说唱 R&B/Soul 驾车	飘飘PPPPP	2015-09-15	https://music.163.co...	只放喜欢的 GFunk ★ Funk ★ TalkBox
5	108154597	HipHop街舞曲	标签： 欧美 说唱 放松	MaeRun	2015-09-18	https://music.163.co...	嘻哈舞曲 经典不断更新
6	111636220	节奏拉《适合编舞的HIP-HOP...	标签： 欧美 说唱 R&B/Soul	Junecleng_1989	2015-09-30	https://music.163.co...	精心整理出来的，节奏感超强，等待大
7	116860650	正牌Bape歌单	标签： 欧美 说唱	_____whatever	2015-10-17	https://music.163.co...	Bape歌单始皇帝
8	121784593	【跟着晃】Trap/Drill/Crunk	标签： 说唱 兴奋	LilPie玩得开	2015-10-31	https://music.163.co...	不会因为封面被和谐吧..... 2019 还是那
9	123624567	Trap说唱 三金用陷阱音乐喂死你	标签： 电子 说唱 另类/独立	崔三金SanJin	2015-11-05	https://music.163.co...	专属于trap的说唱歌单
10	123927037	G.O.O.D. Music	标签： 说唱 欧美	dripnocap	2015-11-06	https://music.163.co...	有时间就会更新 歌曲顺序不分新旧 / 首
11	126338465	健身必备，莱维贝贝	标签： 说唱 运动 欧美	和耀林	2015-11-12	https://music.163.co...	毛霍战士的精神源泉，精挑细选都是自