**YYRating 爬虫程序操作手册**

1. **环境配置**

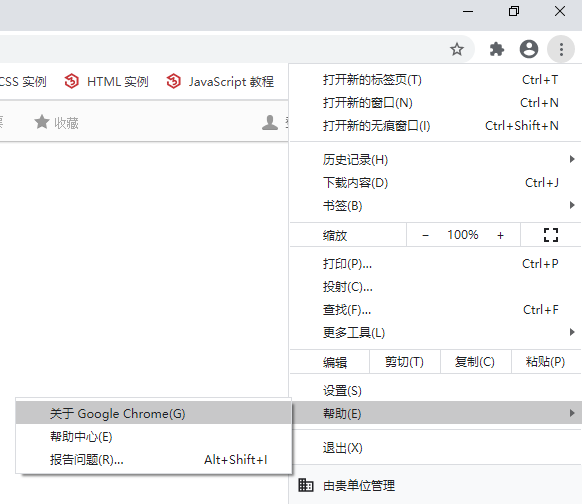
**1.1 配置chromedriver.exe**

（1）在http://npm.taobao.org/mirrors/chromedriver/获取chromedriver

注意：需要下载与谷歌浏览器版本一致的chromedriver版本（向下兼容）

（2）如何获得谷歌浏览器版本：

获得谷歌浏览器版本→浏览器右上角三个点→帮助→关于 Chrome→版本

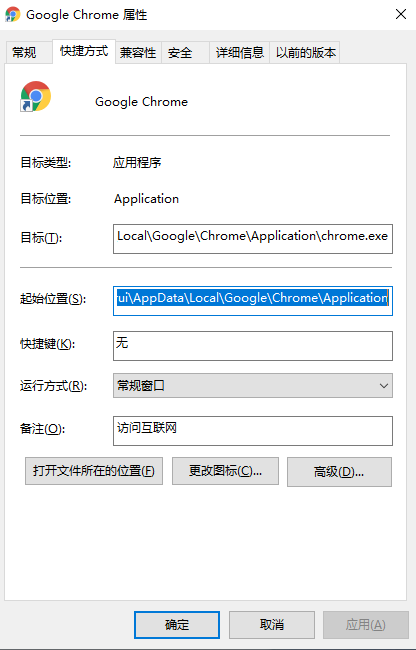


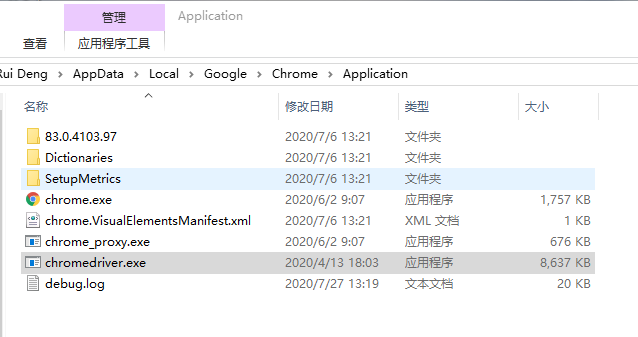


（3）将解压后的chromedriver放到谷歌浏览器的Application目录下

如C:\Users\dengrui\AppData\Local\Google\Chrome\Application

谷歌浏览器属性→打开文件所在的位置→将chromedriver.exe放入

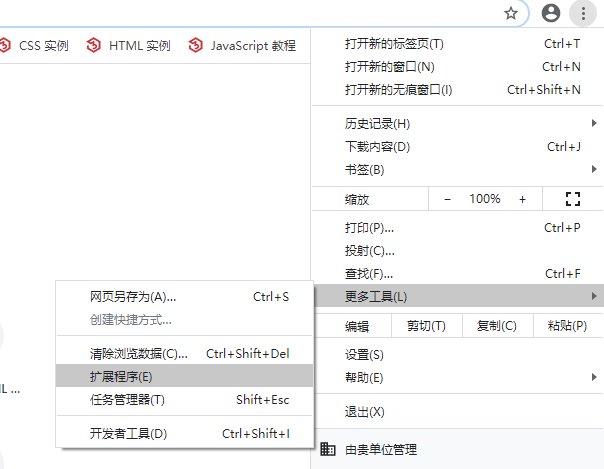


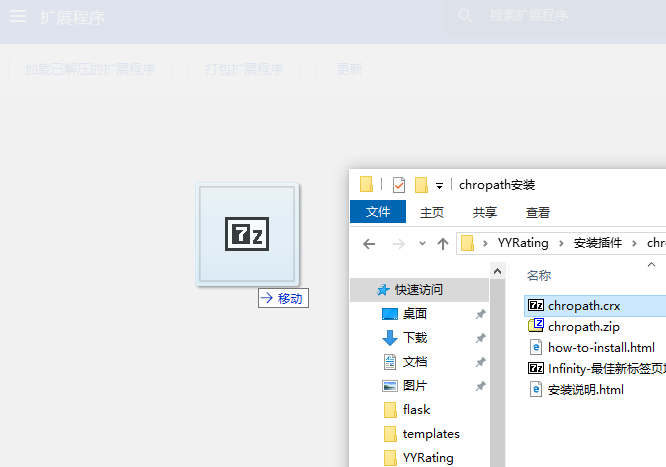


**1.2 安装chropath插件（如果仅运行程序，不需要安装，若需开发其他的爬虫程序，使用selenium包将用到chropath插件）**

（1）将chropath加到谷歌浏览器扩展程序中

打开扩展程序界面→将chropath.crx拖拽进去→添加扩展程序

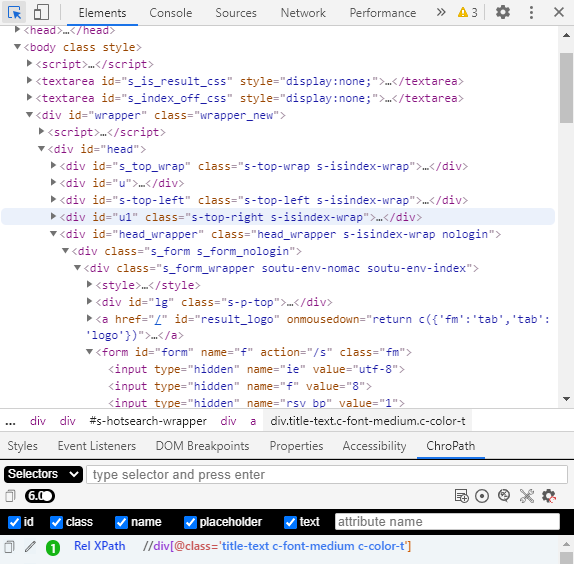






(2) 使用chropath

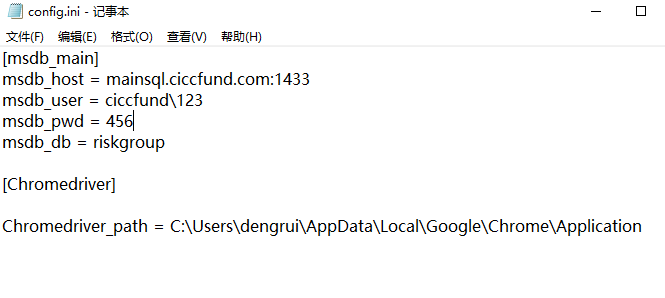
栏目选择Elements，点击最左边的箭头，再选中网页的栏目，最下方的Rel Xpath就是所需要的Xpath



**1.3 配置config**

将数据库连接的配置和Chromedriver的路径添加进config.ini

注意：只需要更改等号右边的内容

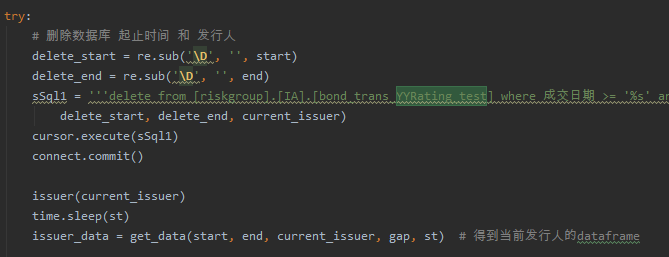
****

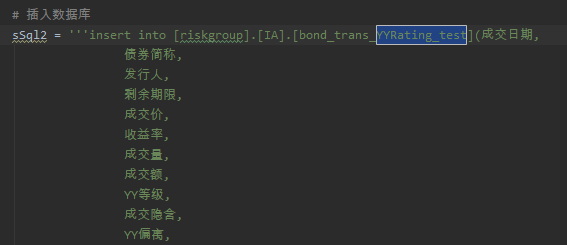
1. **操作运行**

**2.1 事先确认数据库的各种表**

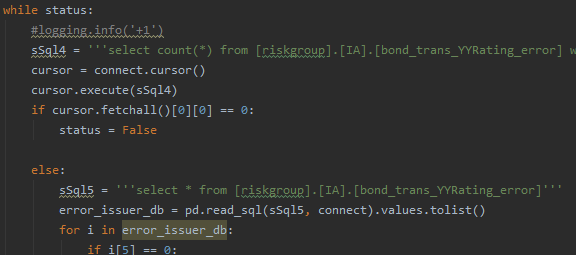
（1）确认查询的发行人表**riskgroup.IA.bond\_issuer\_YYRating**，并将需要查询的发行人插入到该表中

（2）确认插入的表，在程序中为**riskgroup.IA.bond\_trans\_YYRating\_test**





（3）确认纠错的表，在程序中为**riskgroup.IA.bond\_trans\_YYRating\_error**



**2.2 程序运行**

（1）运行的程序是**YYRating\_new.py**

（2）在文件夹的路径下打开cmd输入参数start（开始日期），end（结束日期），gap（间隔）例如：

**YYRating\_new.py --start=2020-01-28 --end=2020-07-28 --gap=30**



注意：start和end的输入均需要输入10位日期格式，gap的取值依据经验应小于等于30，否则会造成错误过多，使纠错过程冗长