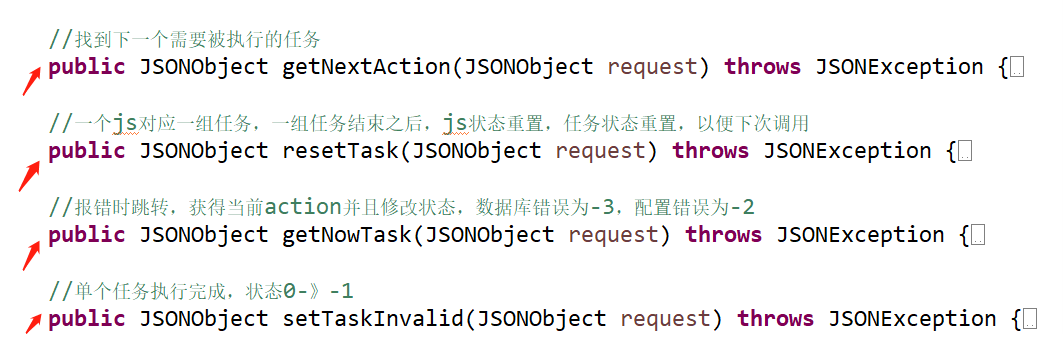
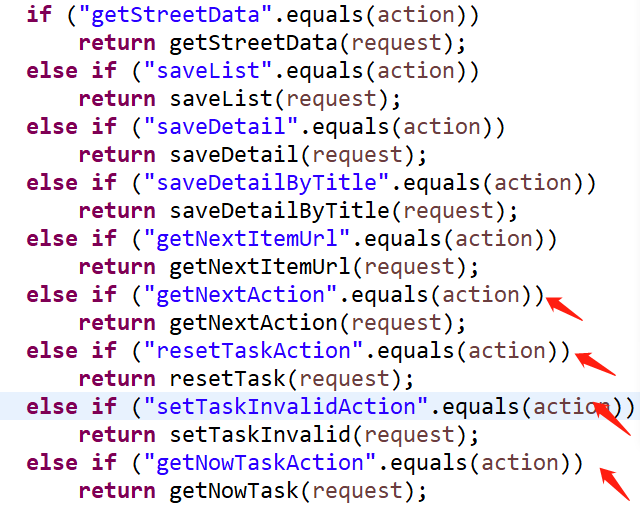
## 修改JAVA代码

在各自的java代码，比如我做的是libSpiderLIBStreetJP，先去重新pull一下项目，确保在C:\Users\ADMINI~1\AppData\Local\Temp\1533864463(1).png，即基类中已经pull到了我们重新修改过的代码，

在BaseLibraryApi中确保已经有如上代码生成，此时回到libSpiderLIBStreetJP（各自的可能是libSpiderLIBStreetCN,libSpiderLIBStreetJP\_02等等）,对action添加如下判断。



上图是我的项目结构，多添加的几行已经标记出来，直接复制就行

else if ( "getNextAction".equals( action ) )

return(getNextAction( request ) );

else if ( "resetTaskAction".equals( action ) )

return(resetTask( request ) );

else if ( "setTaskInvalidAction".equals( action ) )

return(setTaskInvalid( request ) );

else if ( "getNowTaskAction".equals( action ) )

return(getNowTask( request ) );

记得保存，然后检查项目是否有报错，然后重新运行run.bat

## 修改js文件

在spider.load中修改为如下格式,

if (self.updateIndex > 0 && self.dataUrl != null && self.jsUrl != null) {

self.getNextAction();

} else {

self.getNowTaskAndSetTaskInvalid(-2,self.spiderInfo);

setTimeout(function(){

window.location.href = "http://www.baidu.com";

alert("配置错误,检查后刷新页面!\n" + JSON.stringify(self.spiderInfo));

}, 5000);

}

在spider.postData中修改为如下格式

postData: function (params, callback) {

var self = this;

jQuery.post(self.dataUrl, params, function (json) {

if (json.code == 200) {

var data = json.messages.data;

callback(data);

} else {

var error = json.messages.error;

self.getNowTaskAndSetTaskInvalid(-3,"访问服务器故障 {}" + errorCode +"errorCode:" + errorMessage + " errorMessage:");

alert("qwe")

setTimeout(function(){

window.location.href = "http://www.baidu.com";

alert("访问服务器故障:\n" + error.code + "\n" + error.message + "\n请刷新页面");

}, 5000);

}

}, "json");

},

在spider中添加如下代码

getNextAction:function(){

var self = this;

var agentId = self.spiderInfo.sessionId;

var taskId = self.spiderInfo.sourceInfo.id;

self.nextData({

"action": "getNextAction",

"agentId": agentId,

"taskId":taskId

}, function(data) {

console.log(data)

try{

self.workByFunctionName(data.action);

}catch(e){

self.getNowTaskAndSetTaskInvalid(-4,"js code wrong.");

};

setTimeout(function(){

self.setTaskInvalid(data);

window.location.href = "http://www.baidu.com";

}, 5000);

});

},

resetTask:function(){

// 重置task的status为0，并且关闭页面

var self = this;

var agentId = self.spiderInfo.sessionId;

var taskId = self.spiderInfo.sourceInfo.id;

self.nextData({

"action": "resetTaskAction",

"agentId": agentId,

"taskId":taskId

}, function(data) {

self.closeSpider();

});

},

setTaskInvalid:function(data){

// 重置task的status为-1

var self = this;

var agentId = self.spiderInfo.sessionId;

var taskId = self.spiderInfo.sourceInfo.id;

var actionName = data.action;

self.nextData({

"action": "setTaskInvalidAction",

"agentId": agentId,

"taskId":taskId,

"actionName":actionName

});

},

getNowTaskAndSetTaskInvalid:function(errorCode,errorMessage){

var self = this;

var agentId = self.spiderInfo.sessionId;

var taskId = self.spiderInfo.sourceInfo.id;

self.nextData({

"action": "getNowTaskAction",

"agentId": agentId,

"taskId":taskId,

"errorCode":errorCode,

"errorMessage":errorMessage

});

},

workByFunctionName:function(name){

var self = this;

console.log(name)

name = "self." + name + "()";

eval(name);

},

closeSpider:function(){

var self = this;

// 情况storage

self.clearStorage();

//alert("爬取结束");

self.stop();

},

clearStorage:function(){

var self = this;

}

检查各自的spider中是否存在stop,有的话删除添加如下代码,没有的话直接添加如下代码

stop: function () {

window.location = this.spiderInfo.spiderUrl + "?close";

},

同时确保自己爬虫的函数格式为:

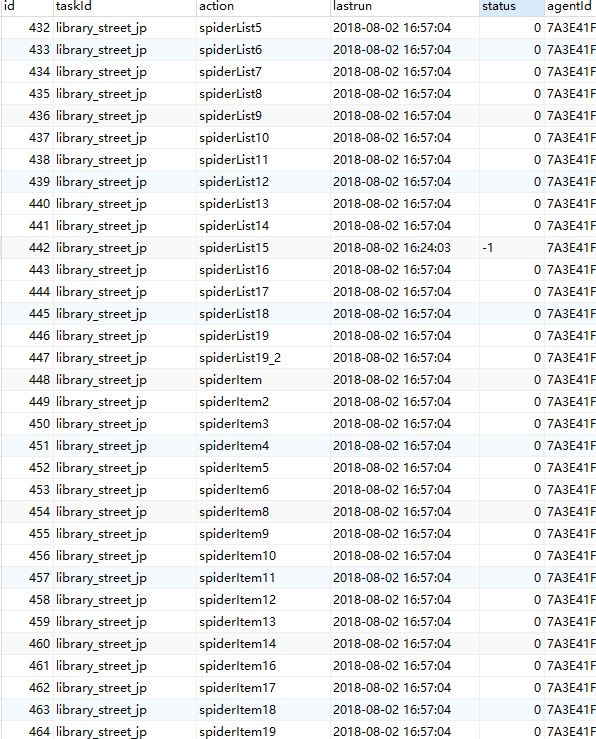
SpiderList或者SpiderItem开头的函数名,如果有部分函数名不对的话先私信江磊或者侯俊杰.

1. 同时回忆一下之前的爬虫是否会一直运行,比如执行spiderListxx,代码会一直运行需要手动停止的请想办法让他停止下来.

## 测试过程

如往常一样运行你的爬虫任务,如果现在你的爬虫任务在执行完spiderListxx之后会自动执行spiderListyy,并且确保没有异常中断或者报错的问题即可.

具体检查方法为，进入数据库xiizii\_dev,找到library\_street\_spider,查看



确保你所有的js代码都在这里出现了（没有出现的话私聊侯俊杰和江磊）.status共有四个状态

0——还未运行

1——正在运行

-1——运行结束

-x——其他负数状态都是表示出现错误

确保位于你js里的所有function都从0变成了-1或者-x，同时最后运行结束的时候这些值又回了0（任务重置等待再次被调用）。

如此你的测试就完成了，私信告诉江磊或者侯俊杰即可。