2017年8月28日12:02:28

运行到：

This is the 323-th loop loading Comments

时报错：

Traceback (most recent call last):

File "E:/\_PythonProjects/Graber\_WarWolf\_WuJing/Grabber2.py", line 66, in <module>

url\_next = 'https://movie.douban.com/subject/26363254/comments' + re.findall(reg, html.text)[0]

IndexError: list index out of range

查看Fiddle得知：

|  |
| --- |
| <script>  var d = [navigator.platform,navigator.userAgent, navigator.vendor].join("|");  window.location.href = "https://sec.douban.com/a?c=e07e80&d="  + d  +"&r=https%3A%2F%2Fmovie.douban.com%2Fsubject%2F26363254%2Fcomments%3Fstart%3D7073&k=hOGf706Kes6osYs9Zinqr5nyhgYex4zwV7m%2F63ylZ9k";  </script> |

Navigator（window.navigator）对象包含有关访问者浏览器的信息。window.navigator 对象在编写时可不使用 window 这个前缀。参见下面示例：

|  |
| --- |
| <!DOCTYPE html>  <html>  <body>  <div id="example"></div>  <script>  txt = "<p>Browser CodeName: " + navigator.appCodeName + "</p>";  txt+= "<p>Browser Name: " + navigator.appName + "</p>";  txt+= "<p>Browser Version: " + navigator.appVersion + "</p>";  txt+= "<p>Cookies Enabled: " + navigator.cookieEnabled + "</p>";  txt+= "<p>Platform: " + navigator.platform + "</p>";  txt+= "<p>User-agent header: " + navigator.userAgent + "</p>";  txt+= "<p>Vendor: " + navigator.vendor + "</p>";  txt+= "<p>User-agent language: " + navigator.systemLanguage + "</p>";  document.getElementById("example").innerHTML=txt;  </script>  </body>  </html> |

|  |
| --- |
| Browser CodeName: Mozilla  Browser Name: Netscape  Browser Version: 5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/60.0.3112.101 Safari/537.36  Cookies Enabled: true  Platform: Win32  User-agent header: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/60.0.3112.101 Safari/537.36  Vendor: Google Inc.  User-agent language: undefined |

**Window Location Href**

location.href 属性返回当前页面的 URL。