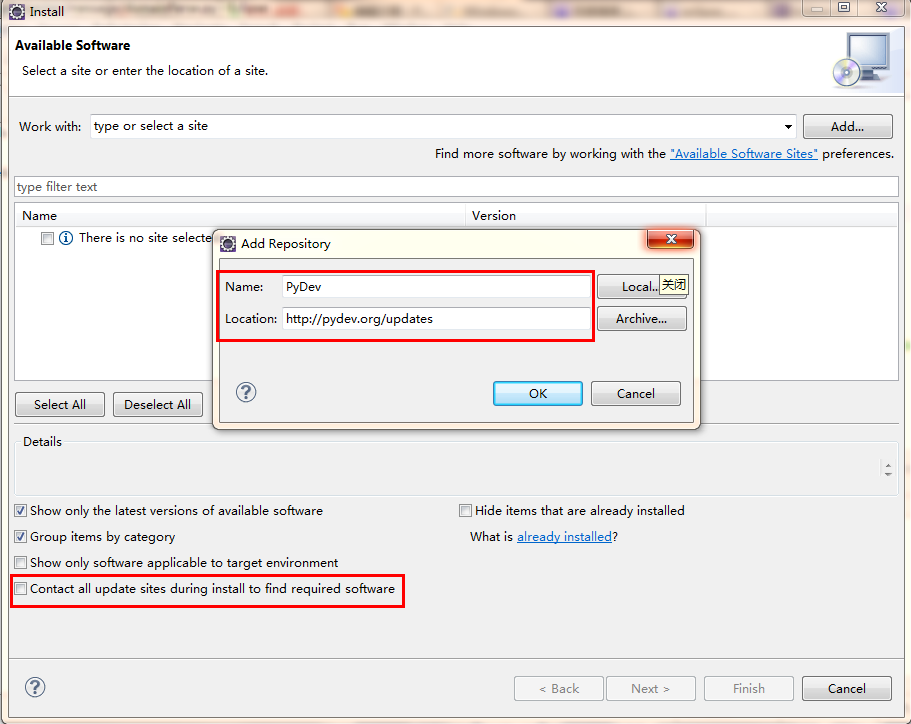
1. Python环境安装

访问<https://www.python.org/getit/>下载最新版本，安装完成后将python安装目录添加到系统变量中。

1. Eclipse配置PyDev插件

打开Eclipse，选择【Help】-【Install New Software】。

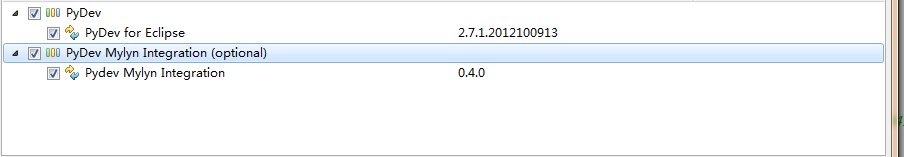
在弹出的对话框中，点击Add按钮，添加新的安装源。



在Location处填写安装源的网址（http://pydev.org/updates），并为之取一个名字填写在Name处，比如我这里写的是PyDev。

把【connect all update sites during install to find required software】的勾选去掉，否则在安装新插件时会联网寻找所有可能的更新站点搜索，导致安装时间不可预估，并可能导致安装失败。

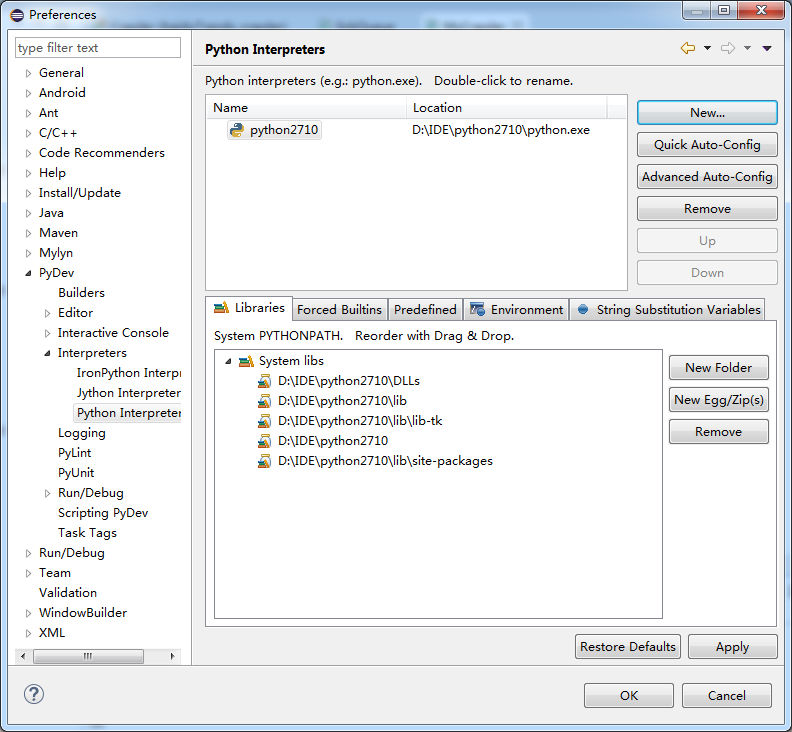
确定后可以看到一个Pending过程，然后得到如下图所示的插件：



勾选后，点击Next进行安装。

安装好PyDev插件后，还并不能正常使用，还需要配置Python的解释器。

打开Eclipse，选择【Window】-【Preferences】，找到【PyDev】，选择其中的【Interpreter – Python】。点击【New】，添加一个系统里已有的Python解释器的位置。确定后会经过短暂的处理，得到它的Libraries、Buildins等。



1. Python第三方类库安装

该爬虫的架构为selenium+PhantomJS，数据存储在MySQL，依赖的第三方类库有两个，分别是MySQLdb和selenium，虽然selenium有调用PhantomJs的API，不过PhantomJS仍需要下载。

**MySQLdb安装**

在<http://sourceforge.net/projects/mysql-python/>下载安装文件，注意适配的Python版本。安装路径使用默认路径。

**Selenium安装**

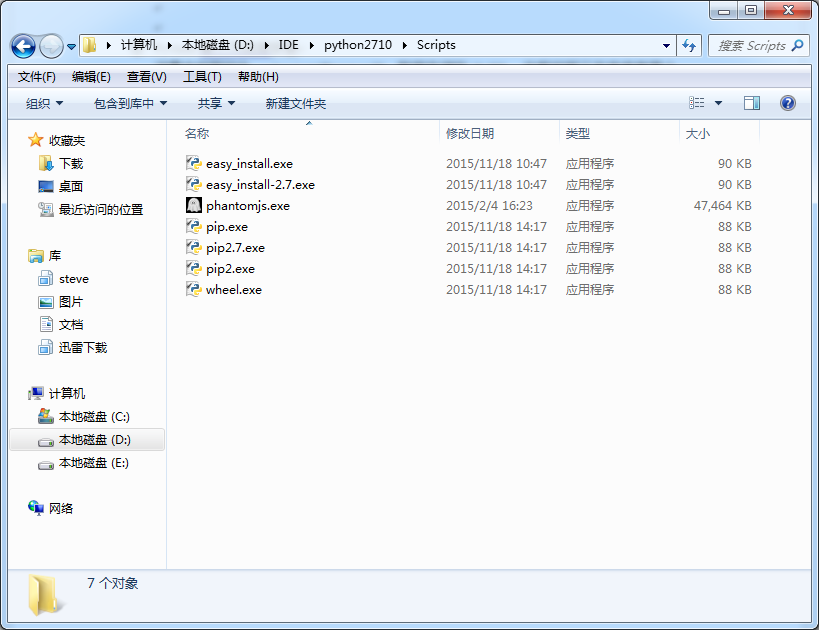
新版Python自带pip，所以可以使用pip在线安装。

直接在python目录里的Scripts目录里运行pip install -U selenium命令即可安装。

**PhantomJS下载**

 Phantom JS是一个服务器端的 JavaScript API 的 WebKit。其支持各种Web标准： DOM 处理, CSS 选择器, JSON, Canvas, 和 SVG。

在<http://phantomjs.org/download.html>下载PhantomJS，解压后将phantomjs.exe这一个文件解压到系统路径所能找到的地方，由于之前我已经将Python下的Scripts目录添加入PATH之中，所以我就直接解压到这个目录。如下：



至此，爬虫运行所需的全部环境配置完成。