XPath与BeautifulSoup的区别

# 原理上的区别

BeautifulSoup是一个第三方的Python库，BeautifulSoup是用python写的。XPath是一门语言，技术。xpath是用c写的。

**BeautifulSoup和XPath的原理不一样**，BeautifulSoup是基于DOM的，会**载入整个文档**，**解析整个DOM树**，**因此时间和内存开销都会大很多**。而lxml只会**局部遍历**。

# 使用区别

XPath 要求一定清楚**文档层次结构**，它通过元素和属性进行导航，可以使用**绝对路径或相对路径**查找，而beautifulsoup 不必清楚文档结构，可以直接找某些标签，简单粗暴。

**打个比方**，如果我们要找这个地址：移动互联网创新园（上海）  
如用正则表达式的话，是这样描述的：它是北边是复旦软件园，门口是高架轻轨线... **如用XPath来描述的话是**：它在 上海市 宝山区 逸仙路 2816号。（从大到小，绝对路径）