

第三课 网络数据的获取与表示

第五节 网页解析模块 (2)

主 讲: Robin

网页解析模块

- 获取节点信息
- 查找节点返回的是Tag对象，获取Tag对象信息

如： `<a class="mnav" href="http://news.baidu.com" name="tj_trnews">新闻</a>`

- Tag.name: 节点标签名称, 'a'
- Tag.attrs: 节点的属性, 字典类型
`{'class': ['mnav'], 'href': 'http://news.baidu.com', 'name': 'tj_trnews'}`
- Tag.text: 获取节点文本, '新闻'

网页解析模块

- 获取其他节点信息
- children: 只返回“孩子”节点
- descendants 返回所有“子孙”节点
- next_siblings 返回下一个“同辈”节点
- previous_siblings 返回上一个“同辈”节点
- parent 返回“父亲”节点

只为遇见明天更优秀的你！