[Scrapy阅读笔记 2](#_Toc23836)

[First Step 2](#_Toc7606)

[初涉Scrapy 2](#_Toc22757)

[简单例子 2](#_Toc21010)

[入门教程 4](#_Toc16718)

[编写第一个爬虫 5](#_Toc4741)

[爬取内容 6](#_Toc29842)

[请求发送流程 6](#_Toc30211)

[Parse回调方法提取内容保存到Item 7](#_Toc16597)

[在Shell中尝试Selector选择器 8](#_Toc24860)

[Response中xpath()/css()/extract()/re()方法由来 9](#_Toc29564)

[提取数据 10](#_Toc17010)

[item的使用 12](#_Toc25669)

[追踪链接(Following links) 重点--双yield写法出处 13](#_Toc12251)

[在命令行保存爬取的数据 14](#_Toc25177)

[命令行工具 15](#_Toc20938)

[默认的Scrapy项目结构 15](#_Toc23209)

[项目控制与操作 16](#_Toc13991)

[可用的工具命令（节选部分重点关注的） 17](#_Toc2633)

[Fetch 17](#_Toc13889)

[Parse 18](#_Toc25778)

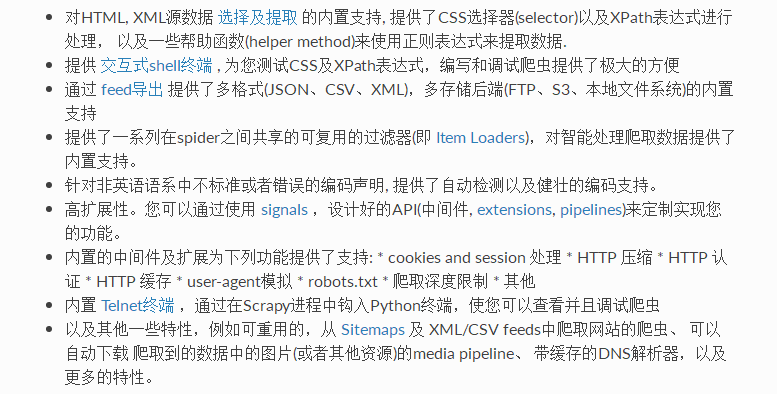
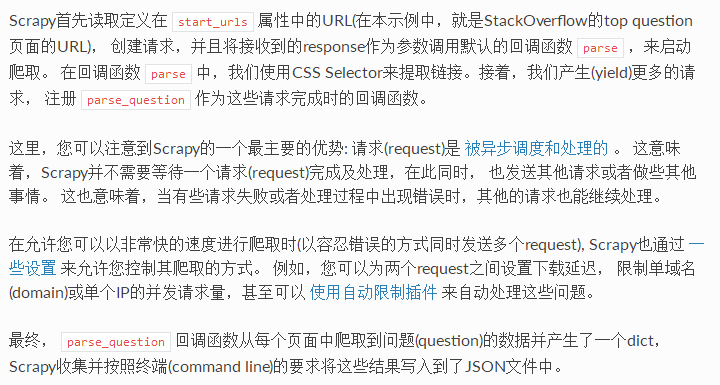
[Spiders 18](#_Toc3458)

# Scrapy阅读笔记

# First Step

## 初涉Scrapy

### 简单例子



## 入门教程



### 编写第一个爬虫



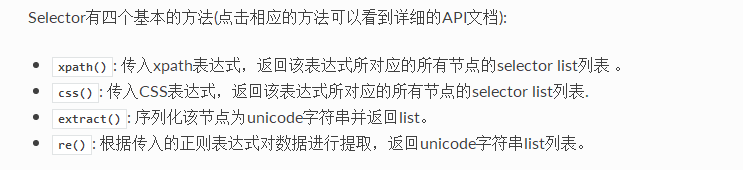
### 爬取内容



### 请求发送流程



### Parse回调方法提取内容保存到Item



### 在Shell中尝试Selector选择器



### Response中xpath()/css()/extract()/re()方法由来



### 提取数据



### item的使用



### 追踪链接(Following links) 重点--双yield写法出处



### 在命令行保存爬取的数据



# 命令行工具

## 默认的Scrapy项目结构





### 项目控制与操作



## 可用的工具命令（节选部分重点关注的）

### Fetch



### Parse



# Spiders

# MongoDB



# PipeLine

## 重复过滤器

