

专栏：001： git 简易教程

E - 爬虫技术 git 教程

1：框架

git教程框架

序号	板块
01	概念介绍篇
02	代码示例篇
03	附属篇

2：概念介绍篇

- 什么是git？
git是一个分布式版本控制软件，由Linus Torvalds创作，最初的目的是为了更好的管理Linux内核开发而设计。
“git”一词源自英国俚语，意思大约是“混账”
- 什么是分布式版本控制？
分布式版本控制又称为分散式版本控制或去中心化版本控制，是一种版本控制的方式。
这种版本控制的特点是：**允许软件开发者可以共同参与一个软件开发专案，但是不必在相同的网络系统下工作**
简单来说**git的版本控制是指：git可以在任何时刻，把文档的更新状态给记录下来，所以可以实现版本更新，回退等一些列操作**
- 分布式 和 集中式版本控制？
集中式版本控制系统，版本库是集中存放在中央服务器。

分布式版本控制系统根本没有“中央服务器”，每个人的电脑上都是一个完整的版本库。

- 为什么学Git?
 - 01: 代码管理
 - 02: 团队协作开发：分支合并，版本更新，协同开发
 - 03: Github号称程序员的“FaceBook”，繁复多样的开源代码

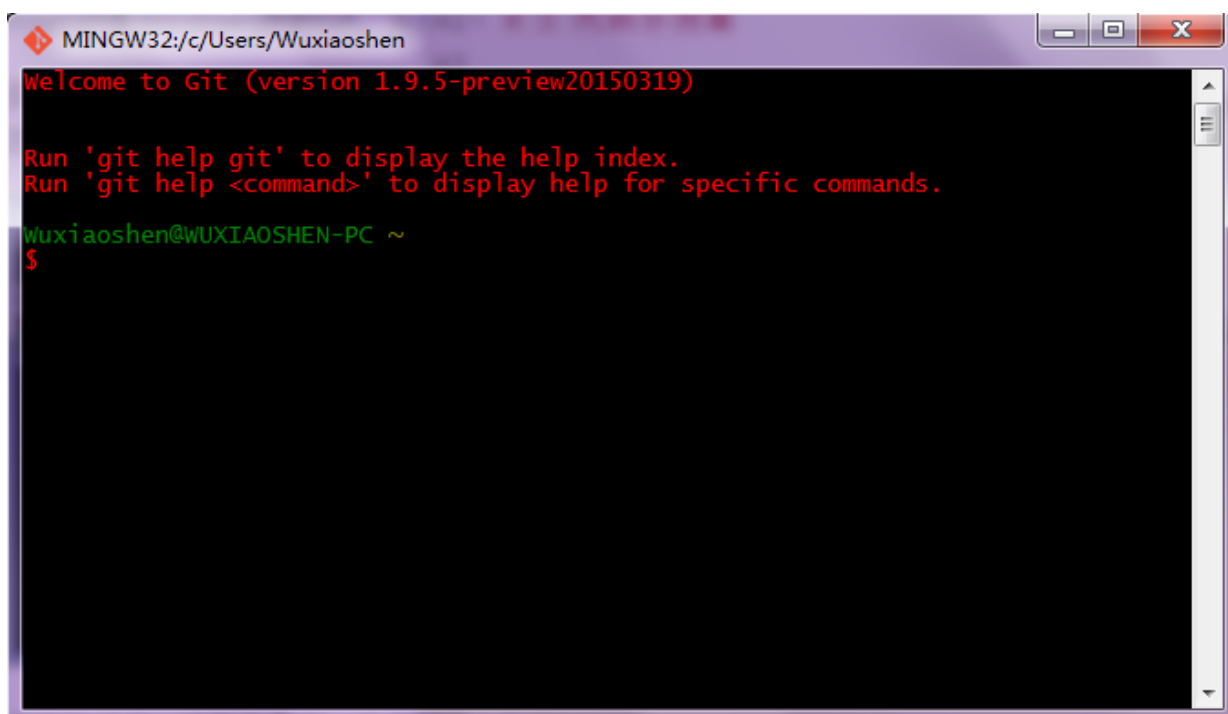
3:代码示例篇

虽然存在各种各样的可视化工具易于理解，但还是推荐命令行操作

以windows平台为例

- 安装git
 - [选择平台进行下载](#)

- 注册Github
 - [Github官网](#)
- 检验是否安装成功



```
MINGW32:/c:/Users/Wuxiaoshen
Welcome to Git (version 1.9.5-preview20150319)

Run 'git help git' to display the help index.
Run 'git help <command>' to display help for specific commands.

Wuxiaoshen@WUXIAOSHEN-PC ~
$
```

界面显示不是默认值，可以在窗口最顶端右键设置

- 初始设置

开始菜单下打开git bash

输入如下命令：

```
git config --global user.name "Your Name"
git config --global user.email "email@example.com"
```

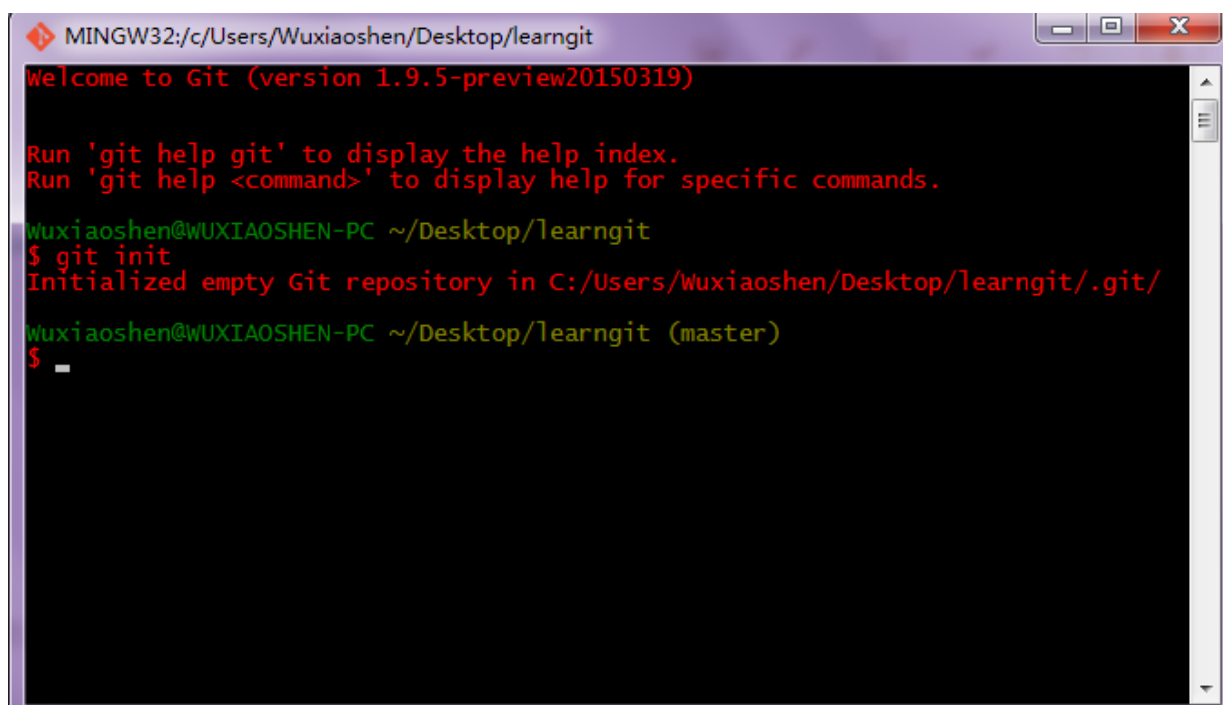
Your Name: 注册github上的用户名

Email: 注册github上的邮箱

- 创建版本库

版本库可以理解为一个文件夹，文件夹下可以放置许多文件

```
01: mkdir learngit      # 新建一个文件夹
02: cd learngit         # 进入文件夹
03: pwd                 # print working directory 显示当前文件夹路径
04: git init            # 创建版本库，此时文件夹下有个隐藏的.git文件夹，
                        这个文件夹用于保存如何git相关的操作记录
```



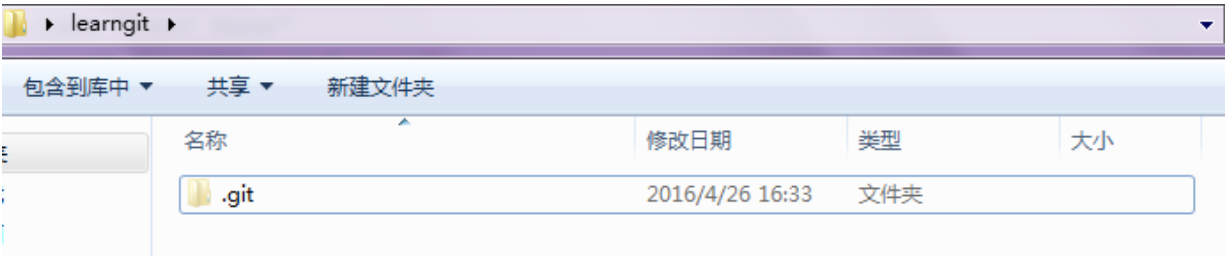
The screenshot shows a Windows command prompt window titled 'MINGW32:/c/Users/Wuxiaoshen/Desktop/learngit'. The window displays the following text:

```
Welcome to Git (version 1.9.5-preview20150319)

Run 'git help git' to display the help index.
Run 'git help <command>' to display help for specific commands.

wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit
$ git init
Initialized empty Git repository in C:/Users/Wuxiaoshen/Desktop/learngit/.git/

wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$
```



- 文件操作

文本编辑器推荐使用：Notepad++

序号	操作	含义	演示
01	git status	查看当前版本状态	<pre>\$ git status On branch master Initial commit Untracked files: (use "git add <file>..." to include in what will be committed) test.txt nothing added to commit but untracked files present (use "git add" to track)</pre>
02	git add	添加文件：没有显示什么就是正确的	<pre>wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learnGit (master) \$ git add test.txt wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learnGit (master) \$</pre>
03	git diff	显示不同：test.txt文件夹进行了修改	<pre>wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learnGit (master) \$ git diff diff --git a/test.txt b/test.txt index 43c6748..405501f 100644 --- a/test.txt +++ b/test.txt @@ -1,2 @@ - how to learn git by yourself? + i can learn by myself. \ No newline at end of file</pre>
04	git commit -m "Add test.txt"	将本地仓库中的文件提交到远程版本库	<pre>wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learnGit (master) \$ git commit -m "Add test.txt" [master (root-commit) 402d5c1] Add test.txt 1 file changed, 1 insertion(+) create mode 100644 test.txt</pre>

以上是对文件的基本操作：
回顾下：

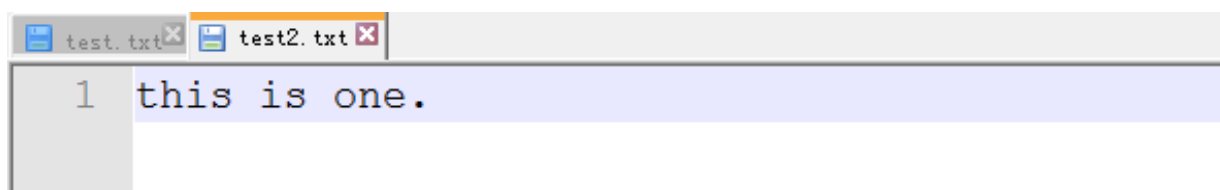
```
git add          # 把文件添加到仓库
git add [file]   # 将指定文件夹添加到仓库
git add .        # 一次性把所有文件添加到仓库
git status       # 查看当前版本库的状态
git diff         # 显示文件的改动状态
git commit -m "添加提交说明"
git log          # 用于记录每次commit 的信息
git reset        # 返回某次的提交
git tag name     # 添加标签
```

- 版本回退操作

在learngit 文件夹下再次添加一个文件：test2.txt

第一次写入内容进行提交

名称	修改日期	类型	大小
.git	2016/4/26 16:59	文件夹	
test.txt	2016/4/26 16:44	文本文档	1 KB
test2.txt	2016/4/26 16:56	文本文档	1 KB



```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git add test2.txt

wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git commit -m "this is one."
[master aee6c47] this is one.
1 file changed, 1 insertion(+)
create mode 100644 test2.txt
```

第二次写入内容进行提交

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git status
On branch master
Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git checkout -- <file>..." to discard changes in working directory)

        modified:   test2.txt

no changes added to commit (use "git add" and/or "git commit -a")
```

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git add test2.txt

wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git commit -m "this is two."
[master d9af71a] this is two.
1 file changed, 2 insertions(+), 1 deletion(-)
```

第三次写入内容进行提交

```
test.txt x test2.txt x
1 this is one.
2 this is two.
3 this is three.
```

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git status
On branch master
Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git checkout -- <file>..." to discard changes in working directory)

        modified:   test2.txt

no changes added to commit (use "git add" and/or "git commit -a")
```

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git diff
diff --git a/test2.txt b/test2.txt
index c30a35f..8569562 100644
--- a/test2.txt
+++ b/test2.txt
@@ -1,2 +1,3 @@
 this is one.
-this is two.
\ No newline at end of file
+this is two.
+this is three.
\ No newline at end of file
```

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git add test2.txt

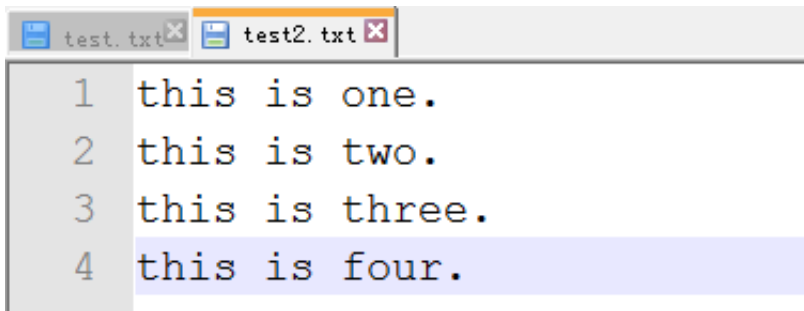
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git commit -m "this is three."
[master dc2cbc1] this is three.
1 file changed, 2 insertions(+), 1 deletion(-)
```

查看文件内容

```
test.txt x test2.txt x
1 this is one.
2 this is two.
3 this is three.
```

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ cat test2.txt
this is one.
this is two.
this is three.
```

再次进行修改文件add但没commit 时，撤销add 操作，返回没被再修改的版本



```
test.txt test2.txt
1 this is one.
2 this is two.
3 this is three.
4 this is four.
```

```
$ git status
On branch master
Changes to be committed:
  (use "git reset HEAD <file>..." to unstage)

        modified:   test2.txt
```

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git reset HEAD test2.txt
Unstaged changes after reset:
M       test2.txt
```

```
$ git status
On branch master
Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git checkout -- <file>..." to discard changes in working d

        modified:   test2.txt

no changes added to commit (use "git add" and/or "git commit -a")
```

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ git checkout -- test2.txt

wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit (master)
$ cat test2.txt
this is one.
this is two.
this is three.
```

- Github上的版本库

登入

```
create new repository
写仓库名: learngit
Create repository
```

克隆至本地

选择本机一个文件夹。

开始菜单: git bash

git clone 网址

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit/new_file (master)
$ git clone git@github.com:wuxiaoxiaoshen/learngit.git
Cloning into 'learngit'...
warning: You appear to have cloned an empty repository.
Checking connectivity... done.
```

显示文件夹

名称	修改日期	类型	大小
 learngit	2016/4/26 17:22	文件夹	

在这个文件夹下添加文件, 进行add,commit等操作, 以后都可以提交到远程仓库

添加一个文件夹: one.txt

内容是: add to remote repository.

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit/new_file/learngit (master)
$ git status
On branch master

Initial commit

Untracked files:
  (use "git add <file>..." to include in what will be committed)

    one.txt

nothing added to commit but untracked files present (use "git add" to track)
```

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit/new_file/learngit (master)
$ git add one.txt

wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit/new_file/learngit (master)
$ git commit -m "one.txt"
[master (root-commit) cf6d645] one.txt
1 file changed, 1 insertion(+)
create mode 100644 one.txt
```

git push origin master :推送至远程仓库

```
wuxiaoshen@WUXIAOSHEN-PC ~/Desktop/learngit/new_file/learngit (master)
$ git push origin master
Counting objects: 3, done.
Writing objects: 100% (3/3), 228 bytes | 0 bytes/s, done.
Total 3 (delta 0), reused 0 (delta 0)
To git@github.com:wuxiaoxiaoshen/learngit.git
 * [new branch]      master -> master
```

返回github仓库learngit显示如下:

wuxiaoxiaoshen / **learngit**

Unwatch

1

Star

0

Fork

0

<> Code

Issues 0

Pull requests 0

Wiki

Pulse

Graphs

Settings

No description or website provided. — Edit

1 commit

1 branch

0 releases

1 contributor

Branch: master

New pull request

New file

Upload files

Find file

SSH

git@github.com:wuxiaoxiaoshen

Download ZIP

wuxiaoxiaoshen one.txt

Latest commit cf6d645 a minute ago

one.txt

one.txt

a minute ago

one.txt的内容

wuxiaoxiaoshen / **learngit**

Unwatch

1

Star

0

Fork

0

<> Code

Issues 0

Pull requests 0

Wiki

Pulse

Graphs

Settings

Branch: master

learngit / one.txt

Find file

Copy path

wuxiaoxiaoshen one.txt

cf6d645 2 minutes ago

1 contributor

1 lines (1 sloc) | 25 Bytes

Raw

Blame

History

1

add to remote repository.

总结下和远程仓库相关的操作：

```
git push origin dev      #提交到远程的dev分支，默认master主分支
git pull origin dev      #拉取代码，比如你在网站上更改了，或者团队更改了代码
git remote add origin github@githu.com:wuxiaoshen/learngit.git # 本地仓库推送至远程
```

4：附属篇

- ssh

SSH公钥是让Git服务器验证你的电脑是被认证的，本地配置SSH避免了每次git要求你输入用户名、密码这些认证信息。
首次需配置秘钥至github网站上
[教程](#)

- .gitignore

目的是：把版本中某些文件并不上传到仓库
文件形式是.gitignore文件，里面写上内容，规定好那些类型的文件可以上传，哪些不可以

标示符	语法	示例
-----	----	----

#	表示备注	这是备注
*	任意通配符	*.pyc 表示忽视所有的后缀名是.pyc的文件
/	文件夹	.idea/:表示忽视整个文件夹
!	表示不忽略	!.gitignore :不忽略这个文件
文件全名	忽略这个文件	001.jpg :忽略001.jpg

5：参考

[廖雪峰git教程](#)
[教程2](#)

关于本人：
初学者，一名半路出家的IT学习者。
Github:[wuxiaoxiaoshen](#)
weibo:[微博](#)

教程预告：爬虫所需要的python 知识 和 正则表达式