

Lecture 14: Bash Shell, Command Line, Grep

LING 1340/2340: Data Science for Linguists

Na-Rae Han

Objectives

- ▶ Finally, shell (bash, zsh)
 - ◆ Running things in command line
 - ◆ Interacting with text files in command line
 - ◆ Regex-based text search using grep

- ▶ Follow up of Lecture 9: Data formats, text file encoding & conversion
 - ◆ <https://naraehan.github.io/Data-Science-for-Linguists-2024/lecture9.pdf>

Bash/Zsh shell

▶ What is a "shell"?

- ◆ [https://en.wikipedia.org/wiki/Shell_\(computing\)](https://en.wikipedia.org/wiki/Shell_(computing))
- ◆ Usually refers to the command-line interface (CLI) as opposed to graphical user interface (GUI).
- ◆ **Bash** is the most common flavor of shell in Unix-like OS.

To find out which shell you're running:

```
echo $SHELL
```

▶ Mac:

- ◆ Mac OS is a Unix-type OS.
- ◆ **Terminal** is a built-in terminal. **Zsh** is the default shell, very similar to bash.

▶ Windows:

- ◆ Not Unix-like OS, so does not come with native bash shell. But we installed "**git bash**": a bash environment for running command-line git.
- ◆ As a bonus, it came with pretty much all of **popular Unix command-line tools**!

Shell introduction, navigating

- ▶ Introducing the shell
 - ◆ <https://swcarpentry.github.io/shell-novice/01-intro.html>
- ▶ Navigating & working with files and directories
 - ◆ <https://swcarpentry.github.io/shell-novice/02-filedir.html>
 - ◆ <https://swcarpentry.github.io/shell-novice/03-create.html>
- ▶ We've been doing some of these already, as part of our git routine. You should know:
 - ◆ `.` `..` `~` `/`
 - ◆ `pwd`
 - ◆ `cd`
 - ◆ `ls`
 - ◆ Command-line history with **↑** and **↓**
 - ◆ Using **TAB** for file name completion
 - ◆ Using **Control+C** to quit

Settling in, customizing

- ▶ You can customize your shell via editing these configuration files:

`.bash_profile`

`.zprofile`

- ▶ In your **home directory**:

- ◆ `your_editor .bash_profile &` ←

- ◆ After adding entries or editing, you should either log back in, or execute `source .bash_profile`

Without **&**, your terminal becomes unusable until you close your editor.

- ▶ Aliasing is the most common customization method:

```
alias calc='/c/windows/system32/calc.exe'
```

```
alias ls='ls -hF --color=tty' ←
```

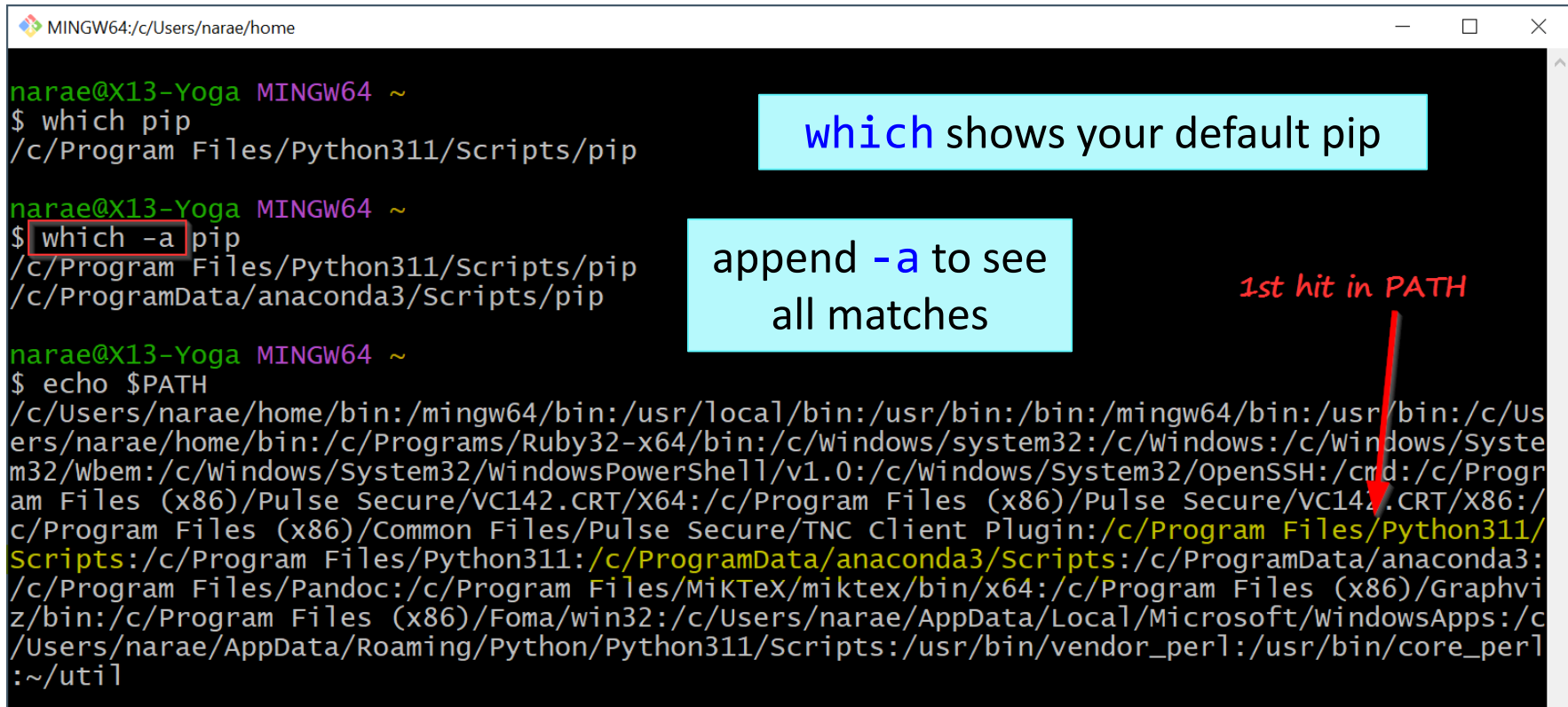
```
alias grep='grep -P --color'
```

← Your favorite shortcuts and command-line options

Mac users: **-G** option for color. You may also have to customize Terminal.

PATH, which, where

- ▶ We have been occasionally using `pip` to install Python libraries. Where is this `pip`? Which `pip` are you using?
 - ◆ Often we end up with multiple distribution versions of Python, meaning multiple `pip` scripts...



```
MINGW64:/c/Users/narae/home
narae@X13-Yoga MINGW64 ~
$ which pip
/c/Program Files/Python311/Scripts/pip

narae@X13-Yoga MINGW64 ~
$ which -a pip
/c/Program Files/Python311/Scripts/pip
/c/ProgramData/anaconda3/Scripts/pip

narae@X13-Yoga MINGW64 ~
$ echo $PATH
/c/Users/narae/home/bin:/mingw64/bin:/usr/local/bin:/usr/bin:/bin:/mingw64/bin:/usr/bin:/c/Us
ers/narae/home/bin:/c/Programs/Ruby32-x64/bin:/c/windows/system32:/c/windows:/c/windows/Syste
m32/Wbem:/c/windows/System32/windowsPowerShell/v1.0:/c/windows/System32/OpenSSH:/cmd:/c/Progr
am Files (x86)/Pulse Secure/VC142.CRT/x64:/c/Program Files (x86)/Pulse Secure/VC142.CRT/x86:/
c/Program Files (x86)/Common Files/Pulse Secure/TNC Client Plugin:/c/Program Files/Python311/
Scripts:/c/Program Files/Python311:/c/ProgramData/anaconda3/Scripts:/c/ProgramData/anaconda3:
/c/Program Files/Pandoc:/c/Program Files/MiKTeX/miktex/bin/x64:/c/Program Files (x86)/Graphvi
z/bin:/c/Program Files (x86)/Foma/win32:/c/Users/narae/AppData/Local/Microsoft/WindowsApps:/c
/Users/narae/AppData/Roaming/Python/Python311/Scripts:/usr/bin/vendor_perl:/usr/bin/core_perl
:~/util
```

`which` shows your default `pip`

append `-a` to see all matches

1st hit in PATH

PATH, which, where

If you want to install tweepy for this copy of python, you can do:

- (1) install from Anaconda navigator
- (2) `/c/ProgramData/anaconda3/Scripts/pip install tweepy`
- (3) cd into `/c/ProgramData/anaconda3/Scripts` directory and then `./pip install tweepy`

```
MINGW64:/c/Users/narae/home
narae@X13-Yoga MINGW64 ~
$ which pip
/c/Program Files/Python311/Scripts/pip

narae@X13-Yoga MINGW64 ~
$ which -a pip
/c/Program Files/Python311/Scripts/pip
/c/ProgramData/anaconda3/Scripts/pip

narae@X13-Yoga MINGW64 ~
$ echo $PATH
/c/Users/narae/home/bin:/mingw64/bin:/usr/local/bin:/usr/bin:/bin:/mingw64/bin:/usr/bin:/c/Us
ers/narae/home/bin:/c/Programs/Ruby32-x64/bin:/c/Windows/system32:/c/Windows:/c/Windows/Syste
m32/Wbem:/c/Windows/System32/WindowsPowerShell/v1.0:/c/Windows/System32/OpenSSH:/cmd:/c/Progr
am Files (x86)/Pulse Secure/VC142.CRT/x64:/c/Program Files (x86)/Pulse Secure/VC142.CRT/x86:/
c/Program Files (x86)/Common Files/Pulse Secure/TNC Client Plugin:/c/Program Files/Python311/
Scripts:/c/Program Files/Python311:/c/ProgramData/anaconda3/Scripts:/c/ProgramData/anaconda3:
/c/Program Files/Pandoc:/c/Program Files/MiKTeX/miktex/bin/x64:/c/Program Files (x86)/Graphvi
z/bin:/c/Program Files (x86)/Foma/win32:/c/Users/narae/AppData/Local/Microsoft/WindowsApps:/c
/Users/narae/AppData/Roaming/Python/Python311/Scripts:/usr/bin/vendor_perl:/usr/bin/core_perl
:~/util
```

1st hit in PATH

Wrapping up

- ▶ Progress report #2 due Friday!
- ▶ Next class
 - ◆ More command line, grep, bash shell scripting
 - ◆ Supercomputing at CRC