[Software Development]

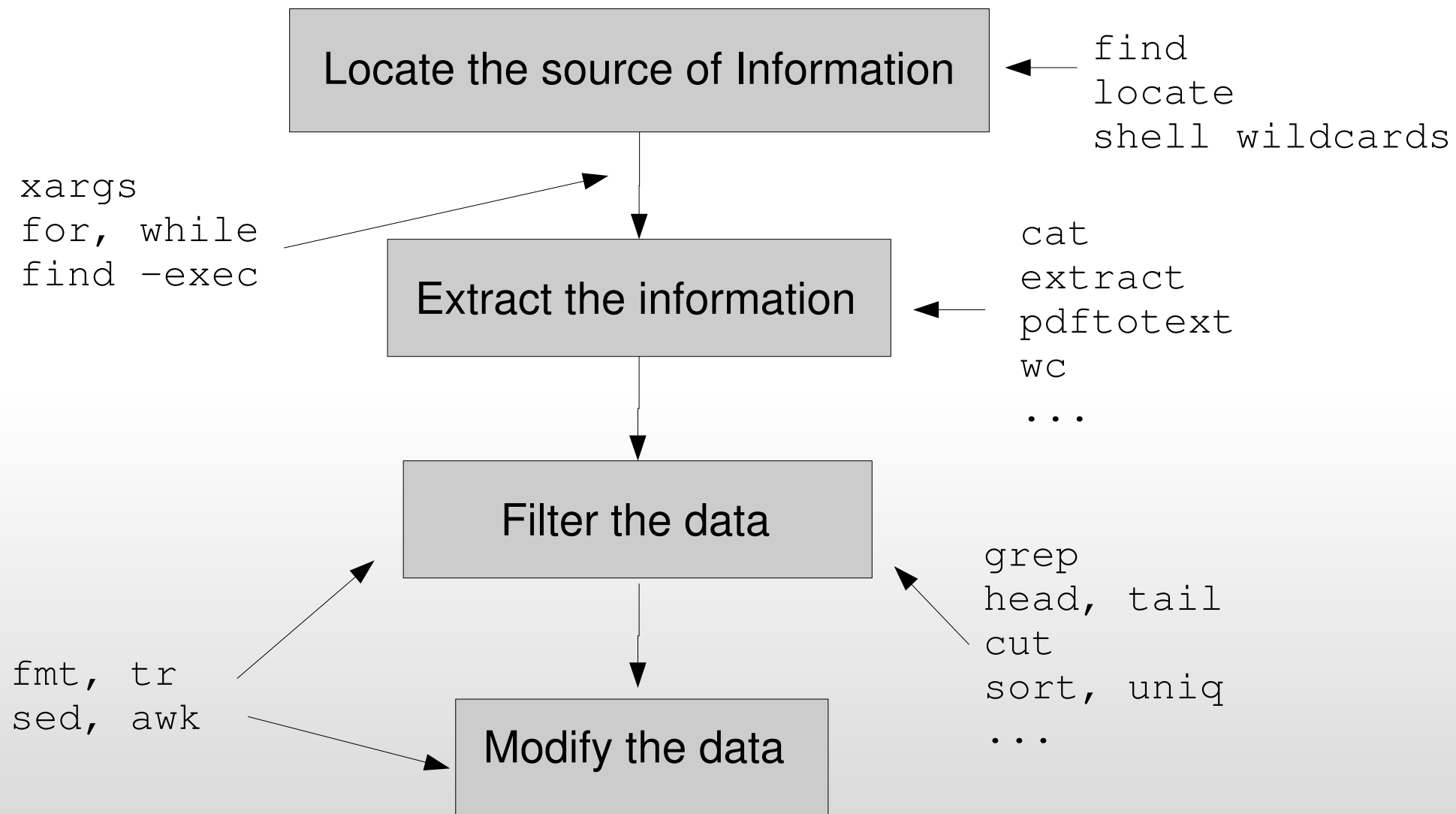# *Linux toolset (part C)*

Davide Balzarotti

Eurecom – Sophia Antipolis, France

- In a text file:

  - Remove everything that is not an alphabetic character

  - Fix duplicated words

  - Print the longest word in a file

- In a text file:

  - Remove everything that is not an alphabetic character

  - Fix duplicated words

  - Print the longest word in a file

- In a C header file:

  - Update the copyright year in the header

  - Replace the comment style from /* .. */ to //

  - Move inline comments to the previous line

- In a text file:

  - Remove everything that is not an alphabetic character

  - Fix duplicated words

  - Print the longest word in a file

- In a C header file:

  - Update the copyright year in the header

  - Replace the comment style from /* .. */ to //

  - Move inline comments to the previous line

- Using "ls -l" print the total file size grouped by file extension

Locate the source of Information ← find
locate
shell wildcards

xargs
for, while
find -exec

Extract the information ← cat
extract
pdftotext
wc
...

Filter the data

fmt, tr
sed, awk

Modify the data

grep
head, tail
cut
sort, uniq
...

The Right Tool For the Right Job

Find files that contains some text	`find	grep`
Number of lines, words, bytes	`wc`
Print certain lines (by number)	`sed`
Print certain lines (by content)	`grep`
Extract patterns	`grep -o`
Remove lines that match a pattern	`grep -v`
Remove certain characters from all lines	`tr`
Remove/substitute a string/pattern	`sed`
Convert to lower/upper case	`tr`
Group lines (w/o frequency)	`sort	uniq -c`
Intersection, union, complements of files	`sort	uniq`

The Right Tool For the Right Job

Remove/keep some columns	`cut, awk`
Repeat a command for each file	`find -exec, find	xargs` `for loop`
Repeat a command for each line	`cat	while loop` `awk`
Repeat a command for each word	`fmt	while, tr	while, awk`
Insert a line	`sed`
Extract part of a file (from ... to)	`sed`
Anything that requires some math	`awk`
Working on record split on multiple lines	`awk`
Modify one or more columns	`awk`

So far...

- Something about the history

  - The origins, the Berkley era, the GNU system...

  - The people: Ken Thompson, Dennis Ritchie, Bill Joy, Richard Stallman, Linus Torvalds...

- Something about the philosophy

  - Write programs that do one thing and do it well

  - Programs are tools and like all good tools, they should be specific in function, but usable for many different purposes

  - Remember the pipe - the power is not in the tools, but in the way you combine them together

- Something about the shell and the command line tools

  - You can go a long way without writing a single line of code

  - Regular expressions can save you in many situation

Scripts

- A way to execute sequences of terminal commands in batch mode

- Text files containing the commands to be executed by a particular interpreter

  - Usually do not follow the compile-link-run cycle, but interpreted or compiled "just in time" (JIT)

- Distinction not so clear nowadays

- Many languages for many different purposes:

  - Shell (bash, C..) - the original, for command scripting

  - Awk – text processing

  - PHP – WWW server side

  - Javascript – WWW client side

  - Perl, Python, Ruby – general purpose languages

A Script under the Hood

```
#!/bin/awk -f

BEGIN { print " - START - "}
{ print $0}
END { print " - DONE -" }
```

A Script under the Hood

```
#!/bin/awk -f

BEGIN { print " - START - "}
{ print $0}
END { print " - DONE -" }
```

If the first line starts with #! the shell considers the rest of the line as the name of the interpreter to be used to run the file

Two ways to run the script

```
> awk -f test.awk
```

```
> chmod a+x test.awk
> ./test.awk
```

A World Full of Script

```
> find /bin /usr/bin/ /usr/local/bin -exec file -b {}
\; | grep -i script | sort | uniq -c
```

A World Full of Script

```
> find /bin /usr/bin/ /usr/local/bin -exec file -b {}
\; | grep -i script | sort | uniq -c

      1 a nickle script text executable
      5 a perl script text executable
      4 a python script text executable
    101 a /usr/bin/perl script text executable
      1 a /usr/bin/perl -t script text executable
      1 a /usr/bin/perl -wl script text executable
     44 a /usr/bin/perl -w script text executable
      1 a /usr/bin/perl -wT script text executable
      2 a /usr/bin/ruby1.8 script text executable
      1 awk script text executable
     35 Bourne-Again shell script text executable
    247 POSIX shell script text executable
     73 python script text executable
      1 setuid a /usr/bin/perl -T script text executable
```