

SED (CH15) / AWK (CH14)

MORE SED SUBSTITUTION

- default: only replaces 1st occurrence of pattern on each line
- can specify to replace others:

- replace all occurrences

```
sed -E 's/pattern_to_match/replace_with/g' fname.txt
```

- replace certain occurrence

```
sed -E 's/pattern_to_match/replace_with/2' fname.txt
```

MORE SED SUBSTITUTION

- combine line range with substitution

```
sed '1,3
```

```
s/pattern_to_match/replace_with/ '
```

```
fname.txt
```

- use matched pattern (or subexpression) in

```
replace_with
```

- \1, \2, ... refer to sub-expressions grouped in parentheses
- & -> entire pattern matched

AWK OVERVIEW

- pattern scanning and processing language
- data driven not procedural
 - describe data to work with
 - tell it what to do once matching data is found
- great for dealing with files with data in columns
- we'll focus on `gawk` the GNU implementation

GAWK PROGRAM

Program has the form:

```
BEGIN { ... initialization awk commands ... }  
{ ... gawk commands to run on each line of the file ... }  
END { ... finalization awk commands ... }
```

- each command has the form: `pattern {
 action }`
- pattern can be regex
- no pattern -> action performed on all lines

FIELDS

- each line made up of fields
- field separator distinguishes fields
 - default = space
 - change value of `FS` to use other
 - or use `-F` option to change
- reference field by `$#` (`$0` is entire line)
- `NF` = number of fields on current line
- `NR` = record number of current line

RUNNING GAWK

- `gawk [options] program [input_files]`
- `gawk -f program_file [input_files]`
- can also create a script telling it to run with gawk instead
 - `#! \bin\gawk -f`
 - needs to be executable