

COMP2041 Week 2 Tutorial

Question 4

Operating system: piece of software that manages the hardware of a computer and provides an interface to the programs that run on the computer.

RegExp

- `^` is the first of line
- `$` is the end of line
- `*` repetition 0 or more
- `+` repetition 1 or more
- `?` 0 or 1
- `|` or
- `[]` inside
- `.` any character
- `\d` digit
- `\D` anything but not digit
- `\s` white space
- `\S` anything besides white space
- `\w` numeric [A-Z0-9]
- `\W` anything but not numeric
- `\b` word boundary
- `\B` not word boundary
- `{n}` n times
- `{n,}` more than n times

Question 5

- `#include` , `#define` , `^#`
- `^[^#]`
- `+$`
- `[BHL P]arry` , `(Barry|Harry|Larry|Parry)`
- `hello.*world`
- `c[ae]l[ae]nd[ae]r`
- `[0-9]+([0-9]+)+`

h. `"[^"]*"*\n"`

Question 6

1. `<` and `>` are unix commands for file input output

add quotes

2. `grep` does not support alteration

use `egrep`

3. tags are case insensitive

use `-i` for case insensitive `egrep -i '<p>|
' /tmp/index.html`

4. there can be spaces between the `<` and `>`

use `-w` `egrep -iw '<p>|
' /tmp/index.html`

5. there can be attributes

`egrep -i '<(p|br)[^>]*>' /tmp/index.html`

Question 8

`[0-9]+\.[0-9]+`

Question 9

`-v` is to exclude, `.` matches any character, so it prints all empty lines in its input

Question 10

`egrep '129\.94\.172\.[1-9]|1[0-9]|2[0-5])$' ips.txt`

Question 11

- a. `:[^:]+:` regex tries to match the first one they can find, so this will first match to the second field

- b. `^[0-9]+(\.[0-9]+)?`

- c. `[^,]+, [^,]+`

- d. `[^,]+,([^,]+)+`

Question 12

- a. `sort q12` , `sort -k2n q12` , `sort -k2nr q12`
- b. `grep 'FL' q12` , `grep '9[0-9] ' q12` , `grep '^[0-9]*[02468] ' q12`
- c. `sort -k2n q12 | head` , `sort -k2nr q12 | head -5`
- d. `cut -d' ' -f3 q12 | sort | uniq -c`

Question 13

- a. `grep 'Ng, Hinry' enrolments | cut -d'|' -f3`
- b. `cut -d'|' -f3 enrolments | sort | uniq | wc -l`
- c. `cut -d'|' -f3 enrolments | sort | uniq -c`
- d. `sort enrolments | uniq | cut -d'|' -f1,2 | uniq -c | egrep -v '^ *1'`