



3ο ΕΡΓΑΣΤΗΡΙΟ

ΣΤΟΧΟΣ

Το φίλτρο grep (egrep) και κανονικές εκφράσεις.

ΑΣΚΗΣΗ 1

(Παρακολουθήστε τα βίντεο «06 - Operating Systems – Regular Expressions» και «07 - Operating Systems – grep» και διαβάστε τις διαφάνειες /OSintro_0[67]_.*\pdf/

Για κάθε κανονική έκφραση που υπάρχει στην πρώτη στήλη του πίνακα που ακολουθεί να χρησιμοποιήσετε την εντολή egrep για να βρείτε ποιοι συνδυασμοί χαρακτήρων από εκείνους που ακολουθούν θα εντοπισθούν.

Κανονική έκφραση	Όνομα αρχείου	1 ^η γραμμή	2 ^η γραμμή	3 ^η γραμμή	4 ^η γραμμή	5 ^η γραμμή
to	file1	hot	To	toad	stool	Top
t.n	file2	tin	Stun	nation	stony	tnuch
a[o-t].	file3	art	Act	task	tort	at

Για να το δοκιμάσετε δημιουργήστε 3 αρχεία (file1, file2, file3). Στο file1 γράψτε (με vi) τις γραμμές:

```
hot
to
toad
stool
Top
```

Και δοκιμάστε την εντολή

```
grep 'to' file1
```

Αντίστοιχα προφανώς και για τα υπόλοιπα αρχεία.



Στην εντολή `grep`, μπορείτε να χρησιμοποιήσετε το option `--color`. τότε χρωματίζεται το σημείο κάθε γραμμής που ταίριαξε στην κανονική έκφραση.

```
asidirop@users: ~  
Linux users.iee.ihu.gr 4.9.0-11-amd64 #1 SMP Debian 4.9.189-3+deb9u2 (2019-11-11) x86_64  
  
Welcome to  
  
International Hellenic University  
Department of Information and Electronic Engineering  
  
We are running Debian GNU/Linux  
  
users.iee.ihu.gr  
(195.251.123.31)  
  
***** NOTE: Please use users.iee.ihu.gr instead of users.itteithe.gr from now on *****  
  
You have new mail.  
Last login: Sat Apr  4 14:37:08 2020 from 46.177.61.104  
asidirop@users:~$ egrep --color 'da.*' /etc/passwd  
daemon:x:1:1:daemon:/usr/sbin:/usr/sbin/nologin  
www-data:x:33:33:www-data:/var/www:/usr/sbin/nologin  
nslcd:x:107:111:nslcd name service LDAP connection daemon,,,:/var/run/nslcd:/bin/false  
asidirop@users:~$
```

ΑΣΚΗΣΗ 2

Συνδεθείτε με `ssh` (ή `putty`) στο `users` εάν δεν το έχετε κάνει ήδη.

1. Δημιουργείτε ένα νέο directory κάτω από το home directory σας, με όνομα της επιλογής σας. Προηγουμένως ελέγξτε αν το όνομα υπάρχει και στη συνέχεια αν έχει δημιουργηθεί.
2. Από τον κατάλογο `~asidirop/public/OSIntro/Lab4_files/` αντιγράψτε στο νέο directory το αρχείο `passwd`, με όνομα `passwd1`. Πριν από την αντιγραφή βεβαιωθείτε ότι το αρχείο υπάρχει. Μετά την αντιγραφή ελέγξτε αν η αντιγραφή έγινε σωστά.
3. Κάθε γραμμή του αρχείου `passwd` έχει την παρακάτω μορφή :
George:xxx:27:107:George_Pap:/home/George:/bin/bash
(7 πεδία με διαχωριστικό την :)
4. Η εντολή `grep` αναζητεί ένα πρότυπο χαρακτήρων μέσα σε ένα αρχείο και επιστρέφει στο αρχείο εξόδου τις γραμμές που περιέχουν το πρότυπο. Βρείτε τη σύνταξη της εντολής (`man grep`) και με την χρήση της :
 - a. Βρείτε μόνον τις γραμμές του `passwd1` που αρχίζουν με τα γράμματα A-M και a-m και βάλτε το αποτέλεσμα σε ένα αρχείο με το όνομα `passwd2`.
 - b. Βρείτε όλους τους χρήστες (τις γραμμές) που χρησιμοποιούν φλοιό `sh` ως προεπιλεγμένο. Το όνομα φλοιού είναι η τελευταία λέξη στο 7^ο πεδίο. Βάλτε το αποτέλεσμα σε ένα νέο αρχείο `passwd3`.



Μπορούμε να ζητήσουμε από το shell, αντί να εκτυπώσει τα αποτελέσματα μιας εντολής στο τερματικό, να «στείλουμε» τα δεδομένα που τυπώνει η εντολή σε ένα αρχείο. Πχ:

```
ls -l > file_test
```

Εκτελείται η εντολή `ls -l` και τα αποτελέσματα δεν θα εμφανιστούν στην οθόνη (τερματικό), αλλά θα αποθηκευτούν στο αρχείο `file_test`. Αν το αρχείο υπάρχει ήδη, τότε επανωγράφεται.

Πχ:

```
grep xyz file1 > file2
```

Θα κάνει αναζήτηση τις γραμμές από το `file1` που περιέχουν το πρότυπο `xyz` και θα τις αποθηκεύσει στο `file2`.

5. Χρησιμοποιείστε την εντολή `wc` για να μετρήσετε πόσες γραμμές αντιστοιχούν στις προηγούμενες δύο περιπτώσεις.

ΑΣΚΗΣΗ 3

Υποθέστε ότι έχετε ένα αρχείο κειμένου που περιέχει τον τηλεφωνικό κατάλογο μιας εταιρείας.

Κάθε γραμμή περιέχει :

{επίθετο} {κόμμα} {κενό διάστημα} {όνομα} {κόμμα} {κενό διάστημα}
{τηλέφωνο}

Έστω το παρακάτω:

```
Francis, John, 5-3871  
Wong, Fred, 4-4123  
Jones, Thomas, 1-4122  
Jones, Tom, 5-4132  
Salazar, Richard, 5-2522  
Salazar, Tom, 3-2521  
Jones, Tommy, 5-2513  
S, R, 6-3433
```

Να εμφανίσετε όλες τις γραμμές (χρησιμοποιώντας εντολές `egrep` και κανονικές εκφράσεις) :

1. Που περιέχουν τηλέφωνο που τελειώνει σε 1.
2. Που περιέχουν τηλέφωνο που αρχίζει από 5 και τελειώνει σε 1.
3. Με όνομα που αρχίζει από J ή W (το όνομα είναι το 2^ο πεδίο στην κάθε γραμμή)
4. Με επίθετο μήκους 4 χαρακτήρων
5. Που δεν περιέχουν τα γράμματα J ή W
6. Με επίθετο που ξεκινά από M έως Z
7. Με επίθετο που ξεκινά από M έως Z και αριθμό τηλεφώνου που τελειώνει με 1 ή 2.
8. Με αριθμό τηλεφώνου που ΔΕΝ τελειώνει σε 22.



ΑΣΚΗΣΗ 4

Υποθέστε ότι έχετε το παρακάτω αρχείο κειμένου με όνομα filep :

```
a
hellohello
goodbye
friend how hello
there how are you how are you
ab
bb
aaa
lll
Parameters
param
```

Να βρείτε τα αποτελέσματα των παρακάτω εντολών και εξηγήστε τα:

1. egrep 'a?' filep
2. egrep '(a|b)+' filep
3. egrep '[ab]+' filep (Ποια είναι η διαφορά με το προηγούμενο;)
4. egrep '.{2}' filep
5. egrep '.{2,}' filep
6. egrep '(.)\1' filep
7. egrep '(.).*\1' filep
8. egrep '(..).*\1' filep
9. egrep '(.{3}).*\1' filep

ΑΣΚΗΣΗ 5

Δημιουργήστε με τον editor vi το παρακάτω αρχείο (με όνομα intro)

```
The Unix operating system was pioneered by Ken
45
AQ
Feb 9, 1999
November 25, 2003
***** position_9 34567 pos_11
bio-informatics course **** position_01
Thompson! and Dennis Ritchie at Bell & Laboratories
in the late 1960s. One of the primary goals in
99 77 88 123
11:45
9:30
00:27
2396-044-777
the design of the UNIX? System* was to create an
environment that promoted **** efficient program
development.
2394-222-789
3:00
08:23
12:00
271.27 tty
2310-555-999
2392-022-111
This is not a ti:me
This is not time: 123:003
```



```
This is not time: 23:003
This is time 23:03 OK

-56.44
123.679

May 29, 1453 phone 2391-333-888
info-question 77.59
Ελληνικά
-.987
```

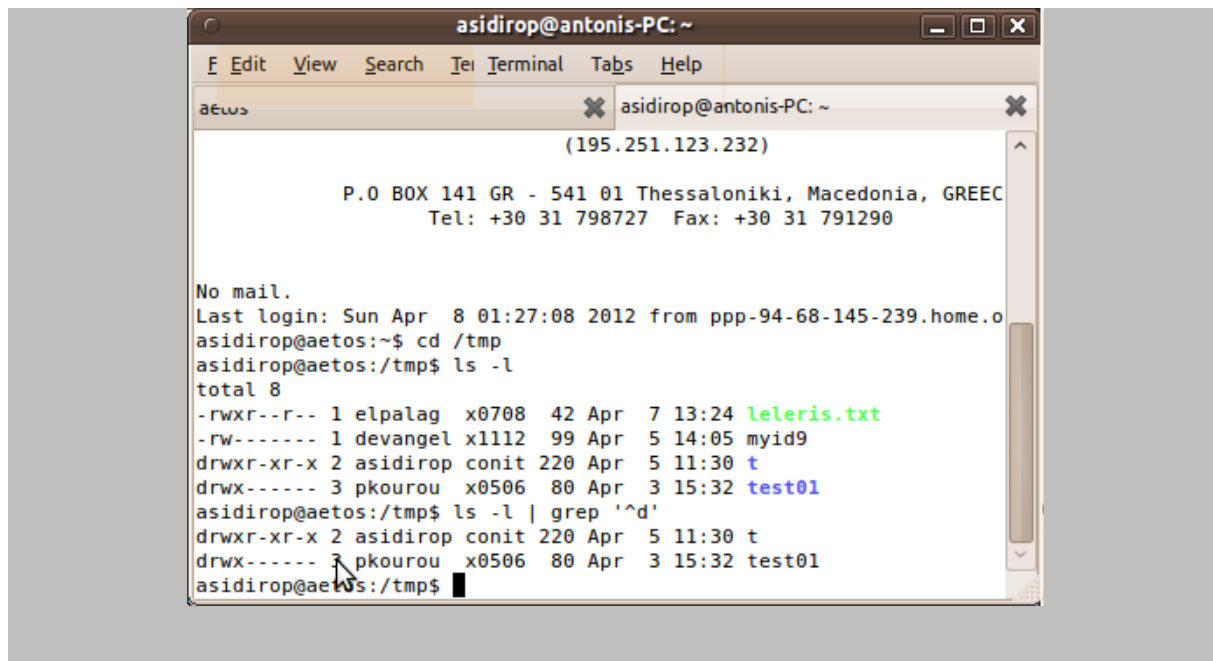
1. Να βρείτε τις γραμμές που περιέχουν ώρα (ψηφία:ψηφία)
2. Να βρείτε τις γραμμές που περιέχουν τηλεφωνικό νούμερο (10 αριθμητικά ψηφία και παύλες)
3. Να βρείτε τις γραμμές που περιέχουν 2 ή περισσότερα συνεχόμενα αστεράκια.
4. Να βρείτε τις γραμμές που ξεκινούν με 2 τουλάχιστον αστεράκια.
5. Να βρείτε τις γραμμές που ξεκινούν με κεφαλαίο γράμμα.
6. Να βρείτε τις γραμμές που ΔΕΝ ξεκινούν με κεφαλαίο γράμμα.
7. Να βρείτε τις γραμμές που περιέχουν 3 ίδιους χαρακτήρες συνεχόμενους (πχ: 333, ***, 999)
8. Να βρείτε τις γραμμές που περιέχουν ΜΟΝΟ αριθμητικά ψηφία (δηλαδή μόνο το 45).
9. Να βρείτε τις γραμμές που περιέχουν ΜΟΝΟ αριθμούς (ακέραιους, δεκαδικούς, αρνητικούς)
10. Να βρείτε τις γραμμές που δεν περιέχουν γράμματα.

ΑΣΚΗΣΗ 6

Μπορούμε βάλουμε 2 ή περισσότερες εντολές να συνεργαστούν μεταξύ τους. Αυτό μπορεί να γίνει χρησιμοποιώντας την διασωλήνωση (pipelining). Πχ:

```
ls -l | grep '^d'
```

Θα εκτελέσει την εντολή `ls -l`, αλλά την έξοδό της δεν θα την εμφανίσει στο τερματικό. Θα δώσει την έξοδο της 1^{ης} εντολής ως είσοδο στην επόμενη εντολή. Η επόμενη εντολή αναζητά τις γραμμές που ξεκινούν με τον χαρακτήρα `d`. Άρα πρακτικά θα εμφανίσουμε μόνο τους καταλόγους.



```
asidirop@antonis-PC: ~
(195.251.123.232)

P.O BOX 141 GR - 541 01 Thessaloniki, Macedonia, GREEC
Tel: +30 31 798727 Fax: +30 31 791290

No mail.
Last login: Sun Apr  8 01:27:08 2012 from ppp-94-68-145-239.home.o
asidirop@aetos:~$ cd /tmp
asidirop@aetos:/tmp$ ls -l
total 8
-rwxr--r-- 1 elpalag  x0708  42 Apr  7 13:24 leleris.txt
-rw----- 1 devangel x1112  99 Apr  5 14:05 myid9
drwxr-xr-x 2 asidirop conit 220 Apr  5 11:30 t
drwx----- 3 pkourou  x0506  80 Apr  3 15:32 test01
asidirop@aetos:/tmp$ ls -l | grep '^d'
drwxr-xr-x 2 asidirop conit 220 Apr  5 11:30 t
drwx----- 3 pkourou  x0506  80 Apr  3 15:32 test01
asidirop@aetos:/tmp$
```

Δημιουργείστε τα αρχεία f1, f2, f3, f4, f5, f6 ,f7 με τα εξής δικαιώματα:

- f1, 757
- f2, 313
- f3, 010
- f4, 642
- f5, 551
- f6, 133
- f7, 111

Με χρήση της εντολής ls -l και egrep να βρείτε τα αρχεία όπου:

1. Το group να έχει τα δικαιώματα r-x.
2. Ο user και οι others έχουν ακριβώς τα ίδια δικαιώματα
3. Ο user, το group και οι others έχουν ακριβώς τα ίδια δικαιώματα
4. Ο user, το group και οι others έχουν τα ίδια δικαιώματα για write και execute
5. Ο user, το group και οι others έχουν τα ίδια δικαιώματα για write
6. Ο user, το group και οι others έχουν τα ίδια δικαιώματα για read και execute
7. Ο user και οι others έχουν τα ίδια δικαιώματα για read και execute

ΑΣΚΗΣΗ 7

Ελέγχοντας το αρχείο ~asidirop/public/OSIntro/Lab4_files/passwd (στο users.it.teithe.gr) και κάνοντας χρήση της εντολής egrep χωρίς να αντιγράψετε το αρχείο (χρησιμοποιείτε την επιλογή -i για να μην γίνεται διάκριση μεταξύ κεφαλαίων και πεζών), βρείτε όλους τους (υποθετικούς) χρήστες του συστήματος όπου:

1. Τουλάχιστον τα 7 πρώτα γράμματα του επιθέτου τους αποτελούν μέρος και του username τους (899 γραμμές)



2. Τουλάχιστον τα 4 πρώτα γράμματα του επιθέτου τους και τα 4 πρώτα γράμματα του ονόματός τους αποτελούν μέρος και του username τους(39 γραμμές)
3. Τουλάχιστον τα 5 πρώτα γράμματα του επιθέτου τους και τα 5 πρώτα γράμματα του ονόματός τους αποτελούν μέρος και του username τους(14 γραμμές)

ΑΣΚΗΣΗ 8

Στο φάκελο ~asidirop/public/OSIntro/Lab5_files υπάρχουν διάφορα αρχεία.

Χωρίς να αλλάξετε τρέχοντα φάκελο εργασίας (παραμένετε δηλαδή στο home directory σας), και χρησιμοποιώντας την εντολή “egrep” με **wildcards** να βρείτε:

1. Σε ποια αρχεία (από όλα) περιέχεται η λέξη file2. Εμφανίστε τις γραμμές (και τα ονόματα αρχείων). Το αποτέλεσμα πρέπει να είναι:

```
/home/staff/asidirop/public/OSIntro/Lab5_files/file2:this is file2.... yes file2
/home/staff/asidirop/public/OSIntro/Lab5_files/file53.txt:grep xyz file1 > file2
/home/staff/asidirop/public/OSIntro/Lab5_files/file53.txt:Θα κάνει αναζήτηση τις γραμμές
από το file1 που περιέχουν το πρότυπο xyz και θα τις αποθηκεύσει στο file2.
```

2. Σε ποια αρχεία περιέχονται γραμμές με μήκος πάνω από 50 χαρακτήρες . Εμφανίστε τις γραμμές (και τα ονόματα αρχείων). (Είναι πολλές οι γραμμές του αποτελέσματος)
3. Σε ποια αρχεία περιέχονται γραμμές με μήκος πάνω από 50 χαρακτήρες. Εμφανίστε μόνο τα ονόματα αρχείων. Ψάξτε το manual της grep για να δείτε πως γίνεται αυτό. Αποτέλεσμα που πρέπει να βγάλετε:

```
/home/staff/asidirop/public/OSIntro/Lab5_files/file2.txt
/home/staff/asidirop/public/OSIntro/Lab5_files/file53.txt
/home/staff/asidirop/public/OSIntro/Lab5_files/file9.txt
/home/staff/asidirop/public/OSIntro/Lab5_files/fileA
```

4. Σε ποια αρχεία περιέχεται κείμενο μήκους τουλάχιστον 5 χαρακτήρων που επαναλαμβάνεται στην ίδια γραμμή (όχι απαραίτητα συνεχόμενα). Εμφανίστε τις γραμμές (και τα ονόματα αρχείων).

```
/home/staff/asidirop/public/OSIntro/Lab5_files/file2:this is file2.... yes file2
/home/staff/asidirop/public/OSIntro/Lab5_files/file2.txt:θα εκτελέσει την εντολή ls -l,
αλλά την έξοδο της δεν θα την εμφανίσει στο τερματικό. Θα δώσει την έξοδο της 1ης εντολής
ως είσοδο στην επόμενη εντολή. Η επόμενη εντολή αναζητά τις γραμμές που ξεκινούν με τον
χαρακτήρα d. Άρα πρακτικά θα εμφανίσουμε μόνο τους καταλόγους.
...
/home/staff/asidirop/public/OSIntro/Lab5_files/file9.txt:ekargiot:x:2936:1025:Kargiotidis
Emmanouil:/home/student/e9697/ekargiot:/bin/bash
staff/asidirop/public/OSIntro/Lab5_files/fileA:ERROR?????
staff/it/asidirop/public/OSIntro/Lab5_files/3.txt OR NOT ERROR ?
```

(Συνολικά 1738 γραμμές)

5. Σε ποια αρχεία με κατάληξη txt περιέχονται (μέσα στο κείμενο) αριθμοί; Εμφανίστε μόνο τα ονόματα αρχείων.

```
/home/staff/asidirop/public/OSIntro/Lab5_files/file2.txt
/home/staff/asidirop/public/OSIntro/Lab5_files/file53.txt
/home/staff/asidirop/public/OSIntro/Lab5_files/file9.txt
```

6. Στο users τρέξτε την εντολή:
getent passwd | grep ~
Ποιο είναι το αποτέλεσμα και γιατί;