

Objectifs : Le but de cette séance est de vous familiariser avec les commandes (filtres) UNIX permettant de réaliser des traitements sur des gros fichiers texte. Seules quelques options offertes par ces commandes seront utilisées : l'aide en ligne (commande **man**) et la pratique seules vous permettront de les utiliser efficacement.

1 Les filtres tail, head et wc

1. Tapez les commandes qui suivent et pour chaque commande que vous ne connaissez pas consultez le manuel en ligne en tapant **man commande**

```
$ mkdir USI ; cd USI
$ cat > fic1
texte 1
très
court
Ctrl-D (^D)
$ ls -la | cat
$ ls -l | cat
$ echo "texte 2" > fic2
$ echo "texte 3" > fic3
$ cat fic1 fic2 fic3 > net
$ cat -n net
$ echo "séparateur" | cat - fic1 fic2
fic3 >> net
$ more net
$ tail -n 2 net
$ tail -c 4 net
# le caractère de fin de ligne
# compte pour un caractère
$ tail +4 net
$ head -1 net
$ head -2 < net
$ wc -l net
$ wc -l < net
$ wc net # remarquez la différence !
$ cat net | wc -l
```

```
$ wc -c net
$ wc -w net
$ cat * | wc
$ wc *
```

2. Extraire (en une ligne de commandes) et affichez sur **stdout** les lignes 4 à 9 incluses du fichier **net**.

2 Le filtre grep

1. Exécutez les commandes suivantes et donnez les significations des options après consultation du manuel en ligne :

```
$ grep "e l" net
$ grep -n "e l" net
$ grep -v sépa net
$ grep -v "sépa très" net
$ grep te..... net
$ grep te.* net
$ grep t net
$ grep ^t net
$ grep "a$" net
$ grep "[a-z]$" < net
$ grep -e "2" -e "3" net
```

2. Faire un lien symbolique sur le fichier **texte**

/net/exemples/ASR2/TP5/french

3. Effectuez les recherches suivantes sur ce dictionnaire à l'aide du filtre **grep** :

- liste des mots se terminant par la chaîne **"cot "**,
- mots commençant par **"av"** et se terminant par **"t"**,
- mots contenant une seule lettre **"a"** et pas de lettre **"c"**,
- mots commençant par une lettre dans l'intervalle **[a-m]** et ne comportant que **cinq** lettres en tout,
- nombre de mots commençant par **"v"**.

3 Les filtres *cut* et *tr*

Dans votre répertoire de travail, faites un lien symbolique vers le fichier

```
/net/exemples/ASR2/TP5/A12
```

et regardez son contenu. Vous devez voir un ensemble de lignes composées de champs. Le format de chaque ligne est le suivant :

```
login:Prénom NOM (groupe):rép.accueil:shell
```

1. Essayez et comprenez (*man tr*, *man cut*) les commandes suivantes :

```
$ cd ~/USI
$ more A12
$ tr ":" " " <A12 | more
$ cut -c0-8 A12 | more
$ cut -c0-2,5-20 A12 | more
$ cut -f 1 A12 -d ":" | more
$ cut -f 1-3 A12 -d ":" | \
  tr ":" " " | more
```

2. À partir du fichier **A12**, affichez une liste comprenant pour chaque utilisateur les trois champs suivants (*séparés par des espaces*) "login" utilisateur, Prénom NOM (groupe), répertoire d'accueil de l'utilisateur. Stockez cette liste dans un fichier **monA12** à l'aide d'une redirection.
3. A partir de ce fichier **monA12**, affichez les utilisateurs de votre promotion et leur nombre.
4. A l'aide du filtre **sort**, affichez une liste triée par ordre alphabétique de tous les "NOMs" d' utilisateurs de votre promotion.

4 Le filtre *sed*

Copiez le fichier **hello.cc** du répertoire **/net/exemples/ASR2/TP5/**

Le filtre **sed** (stream editor) applique des transformations (substitutions, destructions, etc.) à un texte qu'il reçoit en entrée. Essayez :

```
$ sed -e 's/hello/bonjour/' < hello.cc
$ sed -e 's/^ *// ' < hello.cc
```

Les transformations peuvent être

appliquées à des parties du texte

```
$ sed -e '2,$s/hello/bonjour/' \
  < hello.cc
```

```
$ sed -e '1,$d' < hello.cc
```

```
$ sed -e '1,/include/d' < hello.cc
```

Avec cette commande, comment créer un fichier "**g**" constitué des **n-1** premières lignes d'un fichier "**f**" de **n** lignes ?

5 La commande *find*

A l'aide de la commande **find**, écrire une ligne de commande permettant d'afficher le nombre de fichiers suffixés avec ***.log** du répertoire

```
/var/log
```