

TP 2 : Manipulation de chaînes et de fichiers

1 Remarque générale

L'ensemble des manipulations doivent être réalisées dans le shell (sans passer par l'interface graphique). Dès que possible, pensez à utiliser les facilités offertes par le shell (historique, auto-complétion, caractères jokers). Pensez à utiliser le manuel (**man**) pour connaître toutes les options d'une commande.

Pour effectuer vos tests, récupérez les fichiers `germinal.txt`, `germinalSansAccent.txt` et `pascal_pensees.txt` sur ma page web.

2 Option de `grep`

Objectif : `grep` (pour *g/re/p* (*global / regular expression / print*), commande de l'éditeur de texte *ed*) est une des commandes les plus utiles qui permet de chercher directement dans le contenu des fichiers. Cet exercice a pour but de vous faire découvrir les principales options.

Cours : slides 46 et 47.

1. Utilisez la commande `grep` pour trouver toutes les occurrences du mot `ouvrier` dans le fichier `germinal.txt` (remarquez que le logiciel détecte également le mot `ouvriers` au pluriel).
2. Utilisez la commande `grep` pour trouver toutes les occurrences du mot `Belgique` dans le fichier `germinal.txt`. Faites la même recherche mais sans majuscule cette fois "`belgique`". Le résultat est-il le même.
3. Cherchez à nouveau le mot `Belgique` dans `germinal.txt` en ajoutant l'option `-C`. Que se passe-t-il ?
4. Cherchez à nouveau le mot `Belgique` dans `germinal.txt` en ajoutant l'option `-n`. Que se passe-t-il ?
5. Cherchez à nouveau le mot `Belgique` dans `germinal.txt` en ajoutant l'option `-c`. Que se passe-t-il ?
6. Cherchez à le mot `belgique` (sans majuscule) dans `germinal.txt` en ajoutant l'option `-i`. Combien de fois les mots `le` ou `Le` apparaissent-ils dans les deux textes ?
7. Cherchez à nouveau le mot `ouvrier` dans `germinal.txt` en ajoutant l'option `-w`. Que se passe-t-il ?

3 Expressions régulières

Objectif : construire des expressions régulières qui correspondent à la question et tester les avec `grep` sur le fichier `germinal.txt`

Cours : slides 48 à 50

1. Toutes les lignes contenant au moins un chiffre.
2. Toutes les lignes terminant par une majuscule
3. Toutes les lignes commençant par B, E ou Q
4. Tous les mots contenant un r précédé de n'importe quelle lettre majuscule ou minuscule ?
5. Toutes les lignes contenant au moins un mot de 12 lettres.
6. Toutes les lignes contenant au moins deux fois le mot `homme`.
7. Tous les mots parmi : `mange`, `mangea`, `mangeais`, `mangeait`, `mangeaient` (commençant par une majuscule ou une minuscule).

4 Filtrage *

Objectifs : Le but de cet exercice est de vous faire découvrir d'autres utilitaires de Linux et de vous entraîner à les combiner avec des redirections

Cours : slides 51 à 60

Vous aurez besoin des commandes : ps, split, wc, sort, tr, cut, cat, tac, merge, diff, rev. Le manuel et internet sont là pour vous aider à savoir comment fonctionnent ces commandes.

1. Triez les lignes du fichiers `germinal.txt`.
2. Assemblez et triez les fichiers `germinal.txt` et `pascal_pensees.txt`.
3. Cherchez tous les fichiers `.txt` dans le répertoire contenant les répertoires personnels des étudiants de votre filière. Enregistrez le résultat dans le fichier `catalog.txt` et redirigé le résultat vers le fichier spécial `/dev/null`. A quoi correspond ce fichier ?
4. Transformez toutes les lettres minuscules du fichier `pascal_pensees.txt` en lettres majuscules et enregistrez le résultat dans le fichier `PASCAL_PENSEES.txt`.
5. Supprimez tous les espaces du fichier `PASCAL_PENSEES.txt` et enregistrez le résultat dans le fichier `PASCALPENSEES.txt`
6. Découpez le fichier `germinal.txt` en bloc de 100ko.
7. Ré-assemblez les blocs créés au point précédant dans un fichier `germinal2.txt`.
8. Vérifiez que toutes lignes de `germinal` et `germinal2` sont bien égales.
9. Affichez le fichier `germinal` en inversant l'ordre des lignes et en inversant dans chaque ligne l'ordre des caractères.
10. Enregistrez les lignes 45 à 50 du fichier `germinal.txt` dans le fichier `ger45-50.txt`. (commandes `head` et `tail`)
11. Trouvez tous les processus en cours d'exécution appartenant à `root`.
12. Trouvez une commande permettant de n'afficher que les champs contenant le login et le répertoire personnel du fichier `/etc/passwd` (qui contient la liste des utilisateurs du système). (aide : commande `cut`)
13. Extraire les nom de login et UID de `/set/passwd` puis triez les lignes suivant les UID, le tout en une seule commande, vous redirez le tout vers le fichiers `utilisateurs`.
14. Complétez la commande précédente pour remplacer les `:` par des (espaces).
15. Affichez la liste des processus actifs de manière à ce que les colonnes soient séparées par un caractère `.`
16. Triez la liste des processus actif par nom de commande puis si deux commandes portent le même nom par numéro de PID.
17. ** Affichez tous les mots différents contenus dans le fichier avec le nombre d'occurrences de chacun `germinalSansAccent.txt` (aide, le caractère `"\012"` correspond à un saut de ligne, commandes : `cat`, `sort`, `tr`, `uniq`)
18. ** (En utilisant la question précédente) Affichez les 10 mots les plus utilisés du texte `germinal.txt`.
19. ** Créez un fichier `liste` contenant la liste de tous vos fichiers normaux (dont la description détaillée commence par un tiret `-`) avec leur taille (commandes : `cut`, `grep`, `ls` et `tr`).