

TP5 : Textutils, tubes et redirections

- 1.** Afficher les lignes 12 à 15 d'un fichier (commandes `tail` et `head`).
- 2.** Supprimer les lignes 12 à 15 d'un fichier (`tail`, `head`).
- 3.** Modifier un fichier de manière à ce que les lignes 12 à 15 soient triées (`head`, `tail`, `sort` et `cat`).
- 4.** Afficher un fichier en remplaçant les minuscules par des majuscules (`tr`).
- 5.** Afficher un fichier en supprimant les lignes vides (`tr`).
- 6.** Décomposer un texte lu sur l'entrée standard en une suite de mots (sans lignes blanches). Un *mot* est défini comme une suite de caractères alphabétiques minuscules ou majuscules (`tr`).
- 7.** Afficher le dernier mot d'un texte lu sur l'entrée standard (`tr` et `tail`).
- 8.** Lire un texte sur l'entrée standard et donner la liste des mots qui le composent (`tr`, `sort` et `uniq`).
- 9.** Lire un texte sur l'entrée standard et afficher, pour chaque mot, son nombre d'occurrences.
- 10.** Lire un texte sur l'entrée standard et donner les 3 mots les plus utilisés.

Compléments du cours : La commande `tail [-n +nombre] [fichiers]` imprime les fichiers passés en paramètre (ou l'entrée standard si les fichiers ne sont pas spécifiés) en sautant les $n - 1$ premières lignes (par défaut $n = 10$). La version `tail [-n nombre]` permet de ne garder que les n dernières lignes.

Compléments du cours : La commande `head [-n] [fichiers]` écrit sur sa sortie standard les n premières lignes de la concaténation des fichiers indiqués (par défaut $n = 10$).

Compléments du cours : La commande `tr [options] chaîne1 chaîne2` copie l'entrée standard sur la sortie standard en effectuant au passage certaines traductions de caractères. Chaque caractère rencontré appartenant à *chaîne1* est remplacé par le caractère de même rang de *chaîne2* (si *chaîne2* est plus courte que *chaîne1* alors elle est rallongée par répétition du dernier caractère). Des options possibles sont

-d supprimer tous les caractères appartenant à *chaîne1*

-s en sortie, réduire à un seul caractère toute suite de caractères identiques

-c transformer tous les caractères qui ne figurent pas dans *chaîne1*.

Dans l'indication des chaînes, la spécification $[c_1-c_2]$ signifie les caractères compris entre c_1 et c_2 et $\backslash nnn$ signifie le caractère de code nnn en octal. Dernière indication : le code octal $\backslash 012$ désigne le caractère retour-chariot.

Compléments du cours : La commande `uniq [option]` détecte dans son entrée standard les *lignes identiques consécutives* et leur applique un traitement défini par *option*. Les options sont :

-u copier sur la sortie standard les lignes qui ne font pas partie d'un groupe de plusieurs lignes consécutives identiques

-d copier sur la sortie standard une ligne de chaque groupe de plusieurs lignes consécutives identiques

-c copier sur la sortie standard chaque ligne, en la faisant précéder du nombre d'occurrences consécutives identiques

En l'absence d'options, la commande `uniq` se contente de supprimer les lignes superflues (lignes identiques et consécutives).

Compléments du cours : La commande `sort [option]` trie les lignes lues sur son entrée standard et les recopie sur la sortie standard, selon *option* (par défaut, dans l'ordre lexicographique croissant) :

-n dans l'ordre de la valeur numérique du premier mot ;

-r dans l'ordre décroissant.