

TP de Langages de script n° 6 : manipulation de fichiers textuels

Exercice 1 : La disparition

Ecrire un script qui, prenant en entrée un nom de fichier, en crée une “copie” dans laquelle toutes les occurrences du caractère `e` ont été supprimées.

Exercice 2 : Cent mille milliards de poèmes de Raymond Queneau

Le fichier `vers-queneau.txt` contient une séquence de 14 groupements de 10 vers séparés par des lignes vides.

1. Ecrire un script PYTHON chargeant le fichier `vers-queneau.txt` dans un dictionnaire associant aux entiers de 0 à 13, la liste des vers de chaque groupement.
2. Modifier le script pour générer un poème de 14 vers en prenant un vers aléatoirement dans chaque groupement. Utiliser pour cela le module `random` de la bibliothèque standard de PYTHON (<http://docs.python.org/lib/module-random.html>). Ne pas oublier de signer Raymond Queneau à la fin du poème.

Exercice 3 : Une infinité de poèmes de Victor Hugo?

Exercice pouvant compter pour le contrôle continu.

Le fichier `vers-hugo.txt` contient des vers de Victor Hugo, mais ils n’ont pas été classés comme dans l’exercice précédent. Nous allons *essayer* de générer des poèmes par un autre processus.

1. Ecrire un script PYTHON qui charge l’ensemble des vers du fichier dans une liste. On filtrera les entrées suivant le critère suivant (ou un autre qui vous paraîtra plus judicieux): un vers est formé d’au moins 20 caractères et le rapport entre sa taille et le nombre de lettres capitales est supérieur à 5.
2. Modifier le script pour qu’il génère un poème composé de $2n$ vers en utilisant l’algorithme suivant (n étant un paramètre du script):
 - (a) Choisir un vers s_1 au hasard et l’afficher.
 - (b) Tant que $i < 2n$, choisir pour vers s_{i+1} le vers le plus proche de s_i par rapport à la distance:

$$d(s, t) = \frac{20}{1 + d_1(s, t)} + d_2(s, t)$$

où $d_1(s, t)$ désigne le nombre de lettres identiques à la fin de s et de t , et $d_2(s, t)$ la valeur absolue des différences des longueurs de s et de t . Afficher s_{i+1} .

Exercice 4 : Module `re` : échauffement

Pour cet exercice, on va utiliser le module `re` (<http://docs.python.org/lib/module-re.html>).

Dans l’exercice 1 du TP précédent, il était nécessaire de séparer une chaîne de caractères en mots, les mots pouvant être délimités par des espaces ou par tout signe de ponctuation (virgule, point, etc.).

Ecrire en utilisant le module `re` une fonction `liste_mots(phrase)` qui prenne en argument la chaîne `phrase` et renvoie une liste des mots de cette chaîne (ne contenant pas le mot vide).

Exercice 5 : Modifications de l'Education Sentimentale

Exercice pouvant compter pour le contrôle continu.

Nous allons maintenant travailler sur le fichier `educ.txt`.

1. Ce texte est formé de chapitres annoncés par une ligne contenant **Chapitre** N où N est le numéro du chapitre. Ecrire un script chargeant le fichier en mémoire et créant un fichier pour chaque chapitre, intitulé `chapitre-i.txt`.
2. Modifier le script pour qu'il remplace le prénom "Frédéric" par "Gustave".
3. Modifier le script pour qu'il remplace tous les mots commençant par "aim" suivi d'un suffixe par le mot "détest" suivi du même suffixe.
4. Modifier le script pour qu'il extraie les dialogues de chaque chapitre dans un fichier nommé `chapitre-dialogues-i.txt`.