tp_l8dn003_20_02_cor

March 4, 2020

0.0.1 TP n°2 : recherche des anagrammes

1. On part d'un fichier texte qui est chargé en mémoire sous la forme d'une liste de lignes.
2. A partir de ce fichier, il est d'abord demandé de construire un *lexique* : la liste de tous les mots apparaissant dans le fichier. On peut utiliser la fonction `split()` ou essayer d'améliorer le découpage.
3. Ecrire une fonction booléenne `anagramme(x,y)` qui détermine si les deux mots x et y sont des anagrammes l'un de l'autre. On peut utiliser la combinaison de fonctions `sorted(list(x))`.
4. Au moyen d'un double parcours du lexique, afficher toutes les paires de mots qui sont des anagrammes.
5. Construire un dictionnaire indiquant, pour chaque mot du lexique, la liste de ses anagrammes

[2]:
```
def charge_fichier():
    lgns = []
    with open("demo-file-utf8.txt", "r", encoding="utf8") as f:
        for ligne in f:
            lgns.append(ligne.strip())
    return lgns
```

1. Chargement du fichier (on affiche le nombre de lignes)

[3]:
```
liste_lignes = charge_fichier()
print('Le fichier contient %d lignes.' % len(liste_lignes))
```

```
Le fichier contient 5823 lignes.
```

2. Construction du lexique

[4]:
```
# On suppose liste_ligne[] déjà définie
lexique = []
for l in liste_lignes:
    for w in l.split():
        if w not in lexique:
            lexique.append(w)
print('Le lexique formé comprend %d formes.' % len(lexique))
```

```
Le lexique formé comprend 16023 formes.
```

3. Fonction `anagramme` qui utilise l'idiome proposé. La fonction est essayée avec deux valeurs de test.

[5]:
```
def anagramme(a,b):
    return sorted(list(a)) == sorted(list(b))
for x,y in [("maison","bateau"),("lequel","quelle")]:
    print("Les mots '%s' et '%s' %s des anagrammes" % (x,y,"sont" if␣
 ↪anagramme(x,y) else "ne sont pas"))
```

```
Les mots 'maison' et 'bateau' ne sont pas des anagrammes
Les mots 'lequel' et 'quelle' sont des anagrammes
```

4. Parcours de toutes les paires de mots

[7]:
```
# Stratégie la plus brutale
#for w in lexique:
#    for x in lexique:
#        if x != w and anagramme(x,w):
#            print (x,w)

# un peu plus simple: on ne teste les paires x,y que si x < y
for x in lexique[:1500]:
    for y in lexique[:1500]:
        if x < y and anagramme(x,y):
            print (x,y)
```

```
en ne
mon nom
lequel quelle
nos son
courses sources
```

5. Construction du dictionnaire

Non corrigé pour le moment