



MASTER LANGUE ET INFORMATIQUE

Travaux Pratiques n° 10 Manipulation de textes

Les exercices qui vous sont proposés ici ont pour but de connaître les différents formats de représentations des chaînes de caractères et de les manipuler.

1. LECTURE ET RECHERCHE DANS UN TEXTE

Exercice 1 : Ecrire la classe LireToutUnTexte dont le constructeur (de paramètre File) lit un fichier texte (cf. squelette). Ecrire une classe de test avec le fichier « adelaide.txt »

Exercice 2 : Ajouter cette classe, la méthode Chercher qui donne le nombre d'occurrences d'une chaîne de caractères. Ecrire une classe de test avec le fichier « adolphe.txt »

Exercice 3 : Ajouter cette classe, la méthode ChercherVoir qui affiche pour chaque occurrence d'une chaîne de caractères donnée son contexte d'un nombre donnée de mots. Générer la documentation. Ecrire une classe de test avec le fichier « andromaque.txt » et un nombre de mots égal à 3.

```
public class LireToutUnTexte {
    private String sTexte;

    /** lecture d'un fichier texte dans la chaîne sTexte
     * @param ft descripteur du fichier
     */
    public LireToutUnTexte(File ft) {
    }

    /** Recherche du nombre d'occurrences de s dans sTexte
     * @param s chaîne cherchée
     * @return nombre d'occurrences
    }
```

```
*/
public int Chercher(String s) {
}

/** Affichage du contexte des occurrences de s dans sTexte
 * @param s chaîne cherchée
 * @param taille taille du contexte
 */
public void ChercherVoir(String s, int taille) {
}
```

Devoir : Récrire la classe LireToutUnTexte en considérant maintenant qu'elle hérite de la classe BufferedReader. Ecrire une classe de test avec le fichier « armance.txt »

2. EXPRESSIONS REGULIERES

Le but de cette partie est d'utiliser les méthodes Java utilisant les expressions régulières pour étiqueter un texte.

Exercice : Ajouter à la classe LireToutUnTexte, une méthode ChercherER recherchant les occurrences de mots compatibles avec une expressions régulière donnée. Ecrire une classe de test avec le fichier « avare.txt » en recherchant les flexions du verbe parler.

3. ORDRE LEXICOGRAPHIQUE

Le but de cette partie est d'utiliser les méthodes Java permettant de d'ordonner les chaînes de caractères dans différentes langues.

Exercice : Quels sont les différents caractères de la langue esperanto ? Ecrire une classe de test permettant de vérifier la compatibilité de la comparaison standard de Collator avec celle d'un dictionnaire en Esperanto ? (www.esperanto-panorama.net/vortaro/eofr.htm).