

# Travail pratique 1

## Synonymes

Pour le TP1, vous devrez créer un « dictionnaire de synonymes ». Votre application devra être en mesure de fournir une liste des  $n$  meilleurs synonymes pour un mot que l'utilisateur lui fournira. Vous DEVEZ utiliser la méthode décrite dans le document VecteursCooccurrences.pptx.

### Type d'application

Ligne de commande, voir l'image AppelLigneDeCommande.PNG.

### Arguments (respectez l'ordre!)

1. Taille de la fenêtre de cooccurrences.
2. Encodage des textes (ceux que je vous ai fournis sont en utf-8)
3. Chemins des textes à utiliser pour l'entraînement (corpus d'entraînement). Il peut, évidemment y en avoir plus qu'un, mais il pourrait y en avoir 15... Prévoyez cette éventualité.

### Déroulement

Après avoir lancé l'application, l'utilisateur sera invité à fournir un mot pour lequel il veut des synonymes. Il devra taper « q » pour quitter, sinon l'application redemande un mot. L'utilisateur doit fournir le mot, le nombre de résultats à afficher et un nombre qui représente la méthode de score à utiliser :

- 0 : produit scalaire
- 1 : least-squares
- 2 : city-block

Ces informations doivent être données à votre application dans l'ordre décrit et séparées d'un espace.

Vous devez afficher le mot et le score des  $n$  meilleurs résultats, en ordre décroissant du meilleur au moins bon, un par ligne. N'oubliez pas de retirer les stop-words de la liste de résultats!

### À remettre (livrables)

Uniquement les fichiers source.

### Astuce :

Entraînez et testez votre système sur de très, très petits corpus au début (une dizaine de mots) pour voir si tout fonctionne.