

Métaheuristiques
Exercice sur les algorithmes génétiques
Prof. Xavier GANDIBLEUX, Dr. Hab.

Génération automatique de résumé de texte par extraction

1 Description

On souhaite mettre en place une solution logicielle basée sur les algorithmes génétiques qui soit capable de générer automatiquement un résumé de texte par extraction.

2 Activités

1. Prendre en main le code MAIN.JL dans le répertoire NATURALLANGUAGEPROCESSING mis à votre disposition sur Github (<https://github.com/xgandibleux/metaGAstu>). Bien que incomplet, il propose les différents traitements (1) de nettoyage d'un texte donné en vue d'établir un corpus, (2) une fonction d'évaluation d'une phrase autour de trois caractéristiques à l'aide de TF-IDF et Similarité cosinus, (3) des exemples d'appels à ces traitements.
2. Il vous est demandé au cours de cette séance de finaliser la mise en place de l'algorithme générique et de l'évaluer, càd :
 - (a) mettre en place les traitements incomplets (tels que CREATIONPOPULATIONINITIALE, SELECTIONPARENT, CROSSOVER, MUTATION, etc.)
 - (b) proposer un paramétrage du GA
 - (c) rapporter les résultats
 - (d) émettre un commentaire

Un document montrant un exemple d'exécution est disponible sur le dépôt.