

Vous rendrez au plus tard le jour de l'examen une archive compressée contenant

- un rapport au format **PDF** comprenant une description des structures de données et des fonctions utilisées et un commentaire détaillé de vos résultats ;
- les sources de vos programmes contenant impérativement un **makefile** ;
- un fichier **README.txt** contenant une aide pour utiliser vos programmes.

Travail demandé

Implanter l'algorithme d'Aho-Corasick de recherche exacte d'un ensemble de k mots dans un texte.

Utiliser trois méthodes pour représenter l'arbre :

- une matrice de transitions ;
- un tableau de listes d'adjacence ;
- une table de transitions pour la racine et un tableau de listes d'adjacence pour les autres nœuds de l'arbre.

Vos exécutables doivent être nommés **ac-matrice**, **ac-liste**, **ac-mixte** respectivement et prendre deux paramètres, d'abord le nom du fichier qui contient les mots à rechercher puis le nom du fichier qui contient le texte.

Écrire un générateur pseudo-aléatoire de textes. Votre exécutable doit être nommé **genere-texte** et prendre deux paramètres d'abord la longueur du texte à générer puis la taille de l'alphabet. Générer pseudo-aléatoirement des textes de longueur 5 000 000 sur des alphabets de taille 2, 4, 20 et 70.

Écrire un générateur pseudo-aléatoire de mots. Votre exécutable doit être nommé **genere-mots** et prendre quatre paramètres d'abord le nombre de mots à générer puis la longueur minimale et la longueur maximale des mots et enfin la taille de l'alphabet. Pour chacun des alphabets générer pseudo-aléatoirement 3 ensembles de 100 mots de longueur entre 5 et 15, entre 15 et 30 et entre 30 et 60 respectivement.

Pour chacun des textes effectuer la recherche des 3 ensembles avec les 3 implantations de l'algorithme d'Aho-Corasick.

Relever les temps d'exécution de chacune de ces recherches, faire des courbes et commenter.