

## Mini-projet de TP : Bio-Algorithmique

### Table des suffixes

#### Indications

Travail à effectuer **en binôme ou en monôme**  
A envoyer au plus tard le **samedi 13 mai 2023** à l'adresse [recupspace@gmail.com](mailto:recupspace@gmail.com)  
Le travail (compte rendu) doit être tapé dans un éditeur et enregistré au **format pdf**  
Le code source dans un fichier à part  
Nom de fichier = « ProjetTP-nomsEtudiants.pdf »  
L'objet du mail : « Projet-BioALGO »

L'objectif de ce travail est d'implémenter une structure d'index permettant d'accélérer la recherche de motifs, de répétitions, ... dans un texte et d'analyser la complexité spatiale des structures de données utilisées et la complexité temporelle des traitements.

#### Enoncé

**NB :** L'exécution des différents traitements doit être effectuée à travers un menu à choix multiples pour répondre à l'une ou l'autre des questions

Etant donné **un texte T**,

1. Construire sa table des suffixes **TS** et l'afficher
2. Rechercher un motif **M** dans le texte à l'aide de la table des suffixes TS
3. Construire la table **HTR** de T et l'afficher en montrant à chaque fois le préfixe commun entre deux suffixes consécutifs
4. A l'aide des tables **TS** et **HTR**, trouver et afficher
  - le(s) plus long(s) facteur(s) répété(s) dans le texte
  - les facteurs qui se répètent au moins **3** fois.
5. Construire l'inverse **ITS** de la table des suffixes de T et l'afficher
6. Déterminer les plus courts facteurs uniques du texte en construisant la table **LgCandidat**.  
Rappel : un plus court facteur unique est un facteur unique u (qui apparaît une seule fois dans le texte) tel que tous ses facteurs propres ( $\neq u$ ) sont des répétitions.
7. Déterminer les répétitions super-maximales du texte.  
Rappel : une répétition super-maximale est un facteur répété qui n'est facteur propre d'aucune autre répétition.
8. A l'aide des structures TS et HTR, retrouver le plus long facteur commun entre deux textes T1 et T2.

**PS :** Effectuer plusieurs tests en donnant le temps d'exécution sur des textes de tailles variables (100, 200, 600, 1000, 2000, 5000, ...)

Remettre **un rapport** contenant :

1. Introduction et objectifs du travail
2. Description des structures de données utilisées et leur complexité spatiale
3. Pseudo-algorithmes (pour chaque traitement)
4. Les tableaux des tests,
5. Les courbes (ou diagrammes) de variation temporelle
6. Analyse des résultats et conclusions.
7. **Le code source (bien commenté), dans un fichier à part**