

TD 8: Programmation Linéaire et Algorithmes de texte

1 Recherche en espace constant

1.1 Recherche d'un motif auto-maximal

On fixe un ordre total sur l'alphabet Σ , et on note $MaxSuf(w)$ le suffixe maximal de w au sens de l'ordre lexicographique. Le mot w est dit *auto-maximal* si $MaxSuf(w) = w$.

Définition 1 (Période) On dit que $p \in \{1, \dots, |u|\}$ est une période d'un mot $u \in \Sigma^*$ si pour tout $i \in \{1, \dots, |u| - p\}$, $u[i] = u[i + p]$. On note $Period(u)$ la plus petite période de u .

Question 1 : Montrer que $Period(w) = |w| - \pi(|w|)$ où $\pi[i]$ donne la longueur du plus long préfixe propre de $w_1 \dots w_i$ qui est également un suffixe de $w_1 \dots w_i$.

Question 2 : Montrer qu'il existe des mots qui ne sont pas auto-maximaux, peu importe l'ordre choisi sur l'alphabet.

Question 3 : Montrer que si un mot est auto-maximal, alors tous ses préfixes le sont.

On considère l'algorithme suivant :

```

periode_naive(w, j) :=
  pe := 1;
  pour i de 2 à j faire
    si w[i] ≠ w[i - pe] alors pe := i;
  retourner pe;

```

Question 4 : Montrer que, lorsque w est auto-maximal, $periode_naive(w, j)$ calcule bien $Period(w[1, j])$. *Astuce :* Proposer l'invariant de boucle adéquat et le démontrer en utilisant le lemme suivant: pour $p = Period(w[1, i - 1])$, si $w[i] \neq w[i - p]$ alors $w[i] < w[i - p]$ et $Period(w[1, i]) \geq i$.

Question 5 : On suppose que le motif w à rechercher est auto-maximal. Adapter l'algorithme de Morris-Pratt (ci-dessous) pour le faire fonctionner en espace mémoire "constant", plus exactement, le nombre de variables codant des entiers doit rester constant.

```

KMP(P, T, pi) :
  L := empty
  j := 0
  pour i de 1 à |T| faire
    tant que j > 0 et P[j+1] != T[i] faire
      j := pi[j]
    fin tant que
    si P[j+1] = T[i] alors j := j+1
    si j = |P| alors
      AjoutListe(L, i - |P|)
      j := pi[j]
    fin si
  fin pour
  retourner L

```

1.2 Recherche de texte en espace constant

On suppose connue la décomposition $w = u \cdot v$ avec $v = \text{MaxSuf}(w)$, et on cherche les occurrences de w à l'intérieur d'un texte quelconque T .

Question 6 : Si v apparaît avec un décalage de i dans T , montrer que u ne peut apparaître avec un décalage de j dans T , pour tout $j \in \{i - |u| + 1, i - |u| + 2, \dots, i\}$.

Question 7 : (Extra) En déduire un algorithme de recherche de w dans T en temps linéaire et espace constant.

2 LP avec matrices totalement unimodulaires

Une matrice A est dite *totalement unimodulaire* si le déterminant de toute matrice carrée extraite de A appartient à $\{0, 1, -1\}$.

Question 8 : Que peut-on dire des solutions extrémales du système d'équations linéaires donné par $A \cdot x \geq b$, où $A \in \mathbb{Z}^{m \times n}$ est totalement unimodulaire et $b \in \mathbb{Z}^m$? *Astuce :* Utiliser le fait qu'une solution admissible d'un ensemble d'inégalités linéaires dans \mathbb{R}^n est une solution extrême si elle satisfait n inégalités linéairement indépendantes avec égalité.

On considère un graphe orienté pondéré connexe $G = (V, E, c)$ avec deux sommets distingués s et t tels que s n'a pas de prédécesseur et t n'a pas de successeur; c est la fonction de poids de E , $c : E \rightarrow \mathbb{N}^*$. Soit le programme suivant, noté LP, dont les variables sont indicées par les sommets $(\{x_i\}_{i \in V})$ et les arcs $(\{x_{i,j}\}_{(i,j) \in E})$ de G :

$$\begin{aligned} \text{Minimiser } & \sum_{(i,j) \in E} c(i,j) x_{i,j} \text{ tel que} \\ & \forall (i,j) \in E \quad x_{i,j} - x_i + x_j \geq 0 & (1) \\ & x_s - x_t \geq 1 & (2) \\ & \forall (i,j) \in E \quad x_{i,j} \geq 0 \quad \forall i \in V \quad x_i \geq 0 & (3) \end{aligned}$$

Question 9 : Démontrer que le problème admet une solution (réelle).

Il existe donc une solution optimale à LP (LP est réalisable et la fonction objectif est bornée inférieurement).

Question 10 : Démontrer qu'il existe une solution optimale (réelle) qui vérifie:

$$\forall i \in V, x_s \geq x_i \geq x_t \text{ et } \forall (i,j) \in E, x_{i,j} = \max(x_i - x_j, 0)$$

Question 11 : Démontrer qu'il existe une solution optimale (réelle) qui vérifie:

$$\forall i \in V, 1 = x_s \geq x_i \geq x_t = 0 \text{ et } \forall (i,j) \in E, x_{i,j} = \max(x_i - x_j, 0)$$

Question 12 : Démontrer qu'il existe une solution optimale à *valeurs entières* qui vérifie:

$$\forall i \in V, 1 = x_s \geq x_i \geq x_t = 0 \text{ et } \forall (i,j) \in E, x_{i,j} = \max(x_i - x_j, 0)$$

À quel problème de graphe cette solution optimale fournit-elle une réponse?

Astuce 1 : Démontrer que la matrice des contraintes de LP est totalement unimodulaire.

Astuce 2 : D'après la théorie de la PL, LP doit également avoir une solution extrême. Utilisez le fait qu'au moins une solution optimale est une solution extrême.

Question 13 : Calculez le dual de LP. Démontrer que ce dual est équivalent au problème du flot maximal de s vers t dans lequel c est interprété comme la capacité du canal. Que peut-on conclure du théorème de la dualité forte?