

## PLUS LONG SOUS-MOT

Un mot est un synonyme de suite finie de lettres prises dans un ensemble fixé appelé alphabet. Dans ce TP, un mot sera une chaîne de caractères.

Si  $s$  est un mot, alors on appelle sous-mot de  $s$  toute chaîne constituée de lettres prises dans  $s$  dans le même ordre, autrement dit un sous-mot est obtenu à partir du mot initial en supprimant des lettres.

On veut écrire un algorithme qui détermine le plus long sous-mot commun à deux mots donnés.

**1)** Réfléchissez à un algorithme naïf : pourquoi sa complexité calculée en fonction de la longueur des deux mots est-elle un obstacle ?

Une autre méthode s'impose...

On note  $s = s_0s_1 \dots s_{p-1}$  le premier mot de longueur  $p$  et  $t = t_0t_1 \dots t_{q-1}$  le second de longueur  $q$ .

Pour tout  $(i, j) \in \{-1, 0, \dots, p-1\} \times \{-1, 0, \dots, q\}$ , on note  $\ell(i, j)$  la longueur du plus long sous-mot commun à  $s|_i = s_0 \dots s_i$  et  $t|_j = t_0 \dots t_j$ .

**2)** On suppose que  $i$  et  $j$  sont positifs. Si  $m$  est le plus long sous-mot commun à  $s|_i$  et  $t|_j$ , alors plusieurs cas se présentent :

- soit la dernière lettre de  $m$  est différente de  $s_i$ , la dernière de  $s|_i$  : dans ce cas, il est commun à  $s|_{i-1}$  et  $t|_j$
- soit la dernière lettre de  $m$  est égale à  $s_i$  : dans ce cas, on a plusieurs sous-cas :
  - soit la dernière lettre de  $m$  est différente de  $t_j$  : dans ce cas, il est commun à  $s|_i$  et  $t|_{j-1}$
  - soit la dernière lettre de  $m$  est égale à  $t_j$  (donc  $s_i = t_j$ ) : dans ce cas, il est de la forme  $m's_i$  où  $m'$  est un sous-mot commun à  $s|_{i-1}$  et  $t|_{j-1}$

Montrez que le principe de sous-optimalité de Bellmann est valide dans cette situation.

**3)** Donnez une formulation récursive du calcul de  $\ell(i, j)$ .

**4)** Écrivez un algorithme utilisant la programmation dynamique qui calcule la longueur du plus long sous-mot commun à  $s$  et  $t$ .

**5)** Donnez une deuxième version qui calcule en plus un plus long sous-mot commun.

**6)** Donnez une troisième version qui calcule en plus tous les plus longs sous-mot communs.