

MITRO210: Automates et données structurées

Feuille d'exercices 4

Antoine Amarilli

1 Plus long facteur commun

Dans cet exercice, on cherche à identifier un plus long facteur commun entre deux chaînes u et v fournies en entrée, c'est-à-dire un mot w de longueur maximale tel qu'on puisse écrire $u = xwy$ et $v = x'wy'$. On note m et n les longueurs respectives de u et v . Quitte à échanger u et v , on suppose sans perte de généralité que $n \leq m$.

Question 0. Proposer un algorithme naïf pour ce problème. Quelle est sa complexité ?

Question 1. Proposer un algorithme plus astucieux pour ce problème qui atteigne une complexité de $O(mn)$. (Indication : on pourra utiliser la programmation dynamique.)

Question 2. À l'aide des arbres suffixes, proposer un algorithme plus efficace. Quelle est sa complexité ?

Question 3. Peut-on généraliser cette approche à plus de deux chaînes ?

2 Énumération de facteurs

Dans cet exercice, on souhaite trouver efficacement tous les facteurs d'un mot qui appartiennent à un langage.

On fixe un automate A supposé déterministe. Étant donné un mot $u = a_1 \cdots a_n$ de Σ^* , on veut efficacement lister tous les facteurs $1 \leq i \leq j \leq n$ tels que $a_i \cdots a_{j-1}$ soient acceptées par A .

Question 0. Proposer un algorithme naïf pour ce problème en $O(n^3)$. L'améliorer en une complexité de $O(n^2)$. Expliquer en quoi la représentation choisie pour renvoyer les facteurs est importante.

Question 1. Peut-on espérer mieux qu'une complexité de $O(n^2)$ dans le pire cas ?

Question 2. Dans cette question, on fixe le langage a^+ sur l'alphabet $\Sigma = \{a, b, c\}$. Proposer un algorithme pour ce problème en lisant le mot de droite à gauche. Quelle est la complexité de l'algorithme ?

Question 3. On considère maintenant le langage $(aa)^+b$. Étendre l'algorithme de la question précédente.

Question 4. Proposer un algorithme qui énumère les facteurs de la même manière pour un automate déterministe A quelconque.

Question 5. On souhaite à présent résoudre le problème suivant : pour des automates déterministes A_1, A_2, A_3 reconnaissant respectivement les langages L_1, L_2, L_3 , on souhaite énumérer tous les couples (i, j) de positions $1 \leq i \leq j \leq n$ de u telles que l'on a :

- $u_1 \cdots u_{i-1} \in L_1$
- $u_i \cdots u_{i+j-1} \in L_2$
- $u_j \cdots u_n \in L_3$

En quoi ce problème généralise-t-il le problème précédent ? Peut-on résoudre efficacement ce problème ?