

Algorithmique du texte
Contrôle continu — 15 novembre 2011

Documents, calculatrices, mobiles et portables interdits. Durée : 1 h 30. Le barème est donné à titre indicatif.

Exercice I (4 points.)

1) Soit x un mot non vide. Soit p une période de x . Montrez que, pour tout naturel non nul k ,

kp est une période de $x \iff kp \leq |x|$.

2) Quels sont (tous) les mots dont 2 et 5 sont des périodes, mais pas 4. Prouvez ce que vous avancez.

3) Même question avec 3 et 5 périodes, mais pas 4.

4) Même question avec 3 et 5 périodes, mais pas 2.

Exercice II (4 points.)

Pour le mot $x = \text{assassinass}$, dessinez :

1) l'automate implanté par fonction de suppléance et successeur par défaut correspondant à l'algorithme de Morris et de Pratt ;

2) même chose pour l'algorithme de Knuth, de Morris et de Pratt ;

3) l'automate de Simon. Indiquez le décalage des flèches arrière.

Exercice III (5 points.)

Pour l'ensemble de mots $X = \{\text{ara, haras, rat}\}$:

1) Dessinez l'automate de Aho et de Corasick.

2) Donnez la valeur des vecteurs de bits construits dans la phase d'initialisation de l'algorithme PETIT-AUTOMATE.

3) Donnez la valeur du vecteur de bits obtenu après la lecture de chacune des lettres du texte $y = \text{hararat}$ dans la phase de recherche de l'algorithme PETIT-AUTOMATE.

Exercice IV (7 points.)

On cherche à améliorer la phase de recherche de l'algorithme de Knuth, de Morris et de Pratt en s'aidant de la table des décalages de l'algorithme de Horspool : il s'agit de conserver la linéarité du premier algorithme tout en progressant éventuellement plus rapidement sur le texte grâce à la table du second.

On note x le mot cherché, m la longueur de x , y le texte, n la longueur de y , $meil$ -préf la table du meilleur préfixe pour x et d la table des décalages pour x .

Commençons par remarquer que l'expression suivante de l'algorithme de Knuth, de Morris et de Pratt lui donne un aspect d'algorithme de recherche des occurrences de x dans y à fenêtre glissante :

- (1) $j \leftarrow 0$
- (2) $i \leftarrow 0$
- (3) tant que $j \leq n - m$ faire
 - (a) tant que $i < m$ et $x[i] = y[j + i]$ faire
 - (1) $i \leftarrow i + 1$
 - (b) si $i = m$ alors
 - (1) signaler une occurrence de x à la position gauche j dans y
 - (c) $j \leftarrow j + i - \textit{meil-préf}[i]$
 - (d) $i \leftarrow \max\{0, \textit{meil-préf}[i]\}$

1) Proposez-en, en la justifiant, une adaptation intéressante dépendant de $d[y[j + m - 1]]$. (Plusieurs solutions sont possibles.)

2) Illustrez votre adaptation en considérant le cas $x = \textit{assassin}$ (et donc $m = 10$) et $y = \textit{assassassinassociassailliras}$ (et $n = 28$); marquez les comparaisons des lettres de y avec celles de x . Donnez au passage les tables *meil-préf* et *d* de x .