

Algorithmique du texte
Contrôle continu — 17 novembre 2009

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

Exercice I (3 points.)

On rappelle que deux mots x et y sont conjugués lorsqu'il existe deux mots u et v tels que $x = uv$ et $y = vu$.

On note « \bowtie » la relation de conjugaison.

Montrez que la conjugaison est une relation d'équivalence. Autrement dit, montrez que :

- 1 \bowtie est réflexive : $x \bowtie x$, pour tout mot x ;
- 2 \bowtie est symétrique : $x \bowtie y$ entraîne $y \bowtie x$, pour tous mots x et y ;
- 3 \bowtie est transitive : $x \bowtie y$ et $y \bowtie z$ entraîne $x \bowtie z$, pour tous mots x , y et z .

Exercice II (2 points.)

Un mot x est un palindrome s'il est identique à son renversé x^R .

Le bord d'un palindrome est-il un palindrome ? Si oui, montrez-le pour tout mot x ; si non, trouvez un contre exemple.

Exercice III (2 points.)

Quels sont les mots qui admettent 4 et 9 comme période, mais ni 1 ni 7 ? Prouvez ce que vous avancez.

Exercice IV (3 points.)

Dressez :

- 1 la table du bon préfixe pour le mot $x = \text{ananas}$;
- 2 la table du meilleur préfixe pour le mot $x = \text{ananas}$;
- 3 la table du bon suffixe pour le mot $x = \text{sanana}$.

Exercice V (4 points.)

Dessinez DE LA MANIÈRE LA PLUS CLAIRE POSSIBLE la structure utilisée par l'algorithme de Aho et de Corasick pour trouver toutes les occurrences des mots de l'ensemble :

$\{\text{fin}, \text{indice}, \text{défini}, \text{divers}\}$

dans un texte.

Exercice VI (6 points.)

On considère l'algorithme de Horspool de recherche de toutes les occurrences d'un mot x de longueur non nulle m dans un texte y de longueur $n \geq m$. On suppose que la comparaison du facteur $y[j - m + 1..j]$ de y situé dans la fenêtre glissante avec x débute à la position la plus à droite ; si elle doit se prolonger — dans le cas $y[j] = x[m - 1]$ —, aucune hypothèse n'est faite quant à l'ordre des comparaisons des lettres du facteur $y[j - m + 1..j - 1]$ de y avec celles du préfixe $x[0..m - 2]$ de x .

On s'intéresse ici aux mots x dont la dernière lettre, $x[m - 1]$, n'apparaît à nulle autre position dans x .

- 1 Quelle est la valeur du décalage lorsque $y[j] = x[m - 1]$? Que dire de cette valeur dans le cas contraire ?
- 2 Montrez que toute lettre du texte y est comparée au plus 2 fois à une lettre du mot x lors de la recherche.
- 3 Montrez que le nombre de comparaisons entre lettres du texte y et du mot x est inférieur ou égal à n .