

Examen d'Algorithmique du texte

Master 1^{ère} année

Mercredi 7 Mai 2014

Exercice 1. Recherche de motif

1. Calculer la table des bords du mot $u = \text{baabaaab}$.
2. Appliquez l'algorithme de Morris-Pratt pour connaître le nombre d'occurrences de u dans le mot $v = \text{abaabaabaabaabbbbb}$. On détaillera les étapes et on donnera le nombre de comparaisons effectuées.
3. Calculer la table des bords de KMP de u .
4. Appliquez l'algorithme de Boyer-Moore pour connaître le nombre d'occurrences de u dans v . On détaillera les étapes et on donnera le nombre de comparaisons effectuées.
5. Utilisez un des algorithmes vus en cours pour calculer un automate déterministe reconnaissant le langage Σ^*u , avec $u = \text{baabaa}$ et $\Sigma = \{a, b\}$ (Il est possible de dessiner cet automate sans qu'aucune transition n'en intersecte une autre).

Exercice 2. Fouille de données

1. Calculer l'ensemble $X = \text{Fact}(\text{baab})$.
2. Dessiner l'ensemble X muni des ordres partiels \preceq_{prefixe} , \preceq_{suffixe} , \preceq_{facteur} , $\preceq_{\text{sous-mots}}$.
3. Si la relation d'ordre considérée est \preceq_{facteur} , donnez l'idéal du mot aab , noté $\downarrow \text{aab}$ dans l'ensemble X et le filtre du mot ba , noté $\uparrow \text{ba}$.
4. Si la relation d'ordre considérée est $\preceq_{\text{sous-mots}}$, donnez l'ensemble des motifs fréquents de baabaa pour un seuil de fréquence égal à 4.
5. D'après le résultat précédent, le prédicat "u est-il un sous-mot fréquent de v ?" est-il un prédicat monotone ? antimonotone ? Justifiez votre réponse.

Exercice 3. Codage et compression

1. L'ensemble suivant est-il un code $\{011, 110, 1111, 0, 1101\}$? Justifiez votre réponse.
2. Le code $\{0011, 1010, 1101\}$ est-il un code de détection d'erreurs ? de correction d'erreurs ? Si oui, combien d'erreurs peut-il corriger/détecter ? Justifiez votre réponse.
3. Calculez le codage du Huffman du texte suivant :

cddcacaabaadbccacadb

On représentera le code sous la forme d'un arbre, puis son encodage en binaire.

4. Appliquez l'algorithme LZ77 sur le texte

abdcbbbbbcbabdcb

On utilisera une fenêtrre de recherche de taille 10.