Assignment 12

Post Date: 19 Jan 2011   Due Date: 26 Jan 2011, 14:30
You are permitted and encouraged to work in groups of two.

Problem 1: Ukkonen’s Algorithm  

5 Points

Compute a suffix tree for $T = ababbaa \in \Sigma^*$, $\Sigma = \{a, b\}$, with the algorithm of Ukkonen and explain the procedure in detail. Also, indicate the suffix links for all internal vertices of the suffix tree.

Problem 2: Location 

5 Points

Let $sTree(T)$ be a suffix tree for $T \in \Sigma^*$. Assume that the alphabet $\Sigma$ has constant size. Let $P$ be a pattern that occurs in $T$.

Give an algorithm in pseudocode that computes the location of $P$ such that the running time is linear in the size of the traversed vertices.

Problem 3: Longest Common Substring 

6 Points

Give a linear-time algorithm that computes the longest common substring of two strings $S_1$ and $S_2$ over the alphabet $\Sigma = \{a, b, c, \ldots, x, y, z\}$. For example, $S_1 = ararat$ and $S_2 = tiara$, then the longest common substring of $S_1$ and $S_2$ is $ara$.

Problem 4: Algorithm of Boyer & Moore 

4 Points

Show how to compute $s_P^{\text{pref}}(j) = \text{suf}_P(P[j+1, \ldots, m])$ by using the failure function $\pi$ of the Knuth-Morris-Pratt Matcher.