Theory of Computation

Midterm Exam : 23 October 2008

1. Suppose that we have a sequence of n Table-Insert operations on an initially empty table. We want the load factor of a table to be always at least c for some constant $0 < c < 1$. Describe an effective table allocation heuristic and analyze the amortized cost of one Table-Insert operation.

2. Given pattern $P[1..m, 1..m]$ and text $T[1..n, 1..n]$, two-dimensional pattern matching is to find all occurrences of P in T. Describe the Baker-Bird algorithm and its time complexity.

3. Construct the suffix tree for string `aababab#`. Show the suffix tree (and suffix links) in each step of McCreight's construction.

4. Describe the Kärkkäinen-Sanders algorithm for constructing the suffix array. Analyze its time complexity.

5. Show the phrases and output of LZ78 compression when the given string is `ababbaaababaaaabbab`.

6. The off-line minimum problem asks us to maintain a dynamic set T of elements from the domain $\{1, 2, \ldots, n\}$ under the operations Insert and Extract-Min. We are given a sequence S of n Insert and m Extract-Min operations, where each key in $\{1, 2, \ldots, n\}$ is inserted exactly once. We wish to determine which key is returned by each Extract-Min operation. The problem is "off-line" in the sense that we are allowed to process the entire sequence S before determining any of the returned keys.

   (a) Find the returned keys in the following example

   $$4, 8, E, 3, 2, E, 9, 6, E, E, 1, 7, E, 5.$$

   (b) Design an efficient algorithm for the off-line minimum problem using the Union-Find data structure. What is the time complexity of your algorithm? (Hint: The first step is to group the keys into $m+1$ sets as follows $K_1, E, K_2, E, K_3, \ldots, K_m, E, K_{m+1}$. In the example above, $K_1 = \{4, 8\}$, $K_2 = \{3, 2\}$, $K_3 = \{9, 6\}$, $K_4 = \emptyset$, $K_5 = \{1, 7\}$, and $K_6 = \{5\}$.)