**Algorithm Engineering**

**25 June 2013**

**Exercise [rank 4+2]**

Assume that you are given a stream of n items of length L bits each. You wish to process the stream so that at the end, given an item q, you can fast provide an estimate of its number of occurrences.

Discuss the solution, assuming that the distinct items are m < M, but their total length mL > M (so you cannot explicitly store them in memory), and comment on its time complexity and the error probability of the estimate.

**Exercise [rank 4+4]**

Describe the multi-key quicksort and show its execution over the following set of strings (pinco, pallo, box, zoo, mom, dad). Assume that the pivot-char is taken from the first string in the bucket to be sorted.

**Exercise [rank 4+3]**

Given the undirected and weighted graph G={(A,E, 6), (A,F, 2), (A,B, 1), (B,D, 3), (B,F, 5), (C,D, 1), (C,F, 2), (D,F, 1), (E,F, 3)}.

* Compute the MST via the Kruskal’s algorithm
* Compute the MST via the Prim’s algorithm

Show all steps of each algorithm.

**Exercise [rank 6+3]**

Assume that you need to simulate in the external-memory model, with M internal memory and B page size, a parallel algorithm defined to run on P processors. Each parallel step has the form A[ai] = A[ai] + A[bi] , for i=1, 2, … P, where A is an array stored on disk, and ai and bi are the memory cells accessed by the i-th processor in each parallel step.

* Provide an efficient external-memory algorithm which simulates on the single disk this parallel step and compute its I/O-complexity *(hint: use only scan and sort)*.
* Run your algorithm over the following pairs of addresses <ai,bi>, by assuming that P=4 processors: <1,3><2,3><3,4><4,4> and A=[1,3,6,3]

**Exercise [rank \*]**

Define the Steiner tree problem and show how to compute a 2-optimal solution.