HOMEWORK 4

CSC425/520

(1) Prove that in the online bipartite matching problem described in class, if the Shortest Augmenting Path algorithm is used, then if a client node is not matched when it arrives, it will never be matched. Or show a counterexample.

(2) Suppose there is a stream of words in which some words appear exactly once and some appear twice. We have only O(log n) words of storage. At the end of the stream,

The algorithm should output a) the approximate number of words that appear exactly once

b) a word that appears exactly once

For a) and b) explain how to do this with constant probability of success. How good is the approximation for a)?

Extra credit: (a bit complicated)
c) How can we do (a) with $1 - 1/n^c$ probability of success with $O(\log^2 n)$ words of storage?

d) How can we do (b) with $1 - 1/n^c$ probability of success and $O(\log^2 n)$ words of storage?

(3) a) Explain the implementation of the MST alg described in class for map reduce. I.e., Describe the (key, value) pairs which are received and emitted by the mapper and reducer at each step. Your algorithm should take a constant number of phases. The input consists of two pairs for each edge {u,v} in G of the form (u, (u,v)) and (v, (u,v)). The output should consist of the set of pairs ((MST edge), (u,v)) where {u,v} is an edge in the MST for G.

b) For each step, the total data must be $O(N^{1+c/2} \log n)$ words per machine and the size of one (key, value) pair is $O(\log n)$. Explain why this is true of your algorithm.

(4) Explain how the Ben-Or algorithm can fail when N/5 =¡ t¡ N/4 as written in my (corrected) notes in Resources or on page 29 of his paper (in Resources) "B–Byzantine Protocol" .