Exercise IX, Sublinear Algorithms for Big Data Analysis 2024-2025

These exercises are for your own benefit. Feel free to collaborate and share your answers with other students, and solve as many problems as you can. Problems marked (*) are more difficult, but also more rewarding. These problems have been taken from various sources on the Internet, too numerous to cite individually.

1 Prove that any single pass streaming algorithm that tests whether the input graph is connected (and is correct with probability at least 9/10) requires $\Omega(n)$ bits of space.

Solution. We reduce from INDEX. Let $x \in \{0,1\}^n$ be Alice's input in the INDEX problem, and let $i \in [n]$ be Bob's input.

For every $i \in [n]$ Alice creates a pair of vertices $(i, 0)$ and $(i, 1)$. Alice inserts and edge between $(i, 0)$ and $(i, 1)$ for every $i \in [n]$ such that $x_i = 1$. She runs the streaming algorithm on the resulting stream and sends its memory contents to Bob. To learn x_i, Bob adds edges connecting $(i, 0)$ to all other vertices in the graph except $(i, 1)$ (i.e., runs the streaming algorithm on these edges starting with the memory contents that Alice sent him). Note that the resulting graph is connected if and only if $x_i = 1$. □

2 Prove that any single pass streaming algorithm that tests whether the input graph contains a perfect matching (and is correct with probability at least 9/10) requires $\Omega(n^2)$ bits of space.

Solution. We again reduce from INDEX. Let $x \in \{0,1\}^{\binom{[n]}{2}}$ be Alice's input in the INDEX problem, and let $\{i, j\} \in \binom{[n]}{2}$ be Bob's input. We think of entries of x as indexed by edges in the complete graph on n vertices.

Alice inserts an edge $\{a, b\}$ into the stream for every $\{a, b\} \in \binom{[n]}{2}$ such that $x_{\{a,b\}} = 1$. Bob adds a new vertex a' for every $a \in [n]$ except $a = i$ or $a = j$ and inserts an edge $\{a, a'\}$ for every such a. Note that the graph contains a perfect matching if and only if $x_{\{i,j\}} = 1$. Thus, an algorithm for the perfect matching problem can be used to solve INDEX on vectors of length $\binom{n}{2}$, and therefore the space complexity of perfect matching is $\Omega(n^2)$, as required. □