Advanced Algorithms 11/01, 2019

Exercise 07

Lecturer: Mohsen Ghaffari *Teaching Assistant: Saeed Ilchi*

NOTE: All problems discuss a data stream of length m of elements in $\{1, 2, \ldots, n\}$.

1 Majority [Recommended]

In the class, we discussed the deterministic algorithm of Boyer and Moore which has space complexity $O(\log n)$ and outputs an element such that if the stream has a majority, the output is equal to this majority. Prove that any deterministic algorithm that is able to decide whether a stream has a majority element or not needs space $\Omega(n)$ bits.

Hint 1:

Alice and Bob each get a subset of $\{1, 2, \ldots, n\}$ and want to know whether these sets have non-empty intersection or not. They want to solve the problem by the least amount of communication and only Alice is allowed to send a message to Bob. How many bits are necessary in order to Bob find out the answer by looking at his subset and Alice's message.

Hint 2:

Suppose there is a streaming algorithm that decides whether a stream has a majority element or not with space k bits. How many bits are sufficient to communicate by Alice, as a function of k, to solve the problem in Hint 1?

2 Frequent Elements [Recommended]

Extend the majority algorithm of Boyer and Moore, which we discussed in the class, to an algorithm with space complexity of $O(k \log n)$ bits that outputs k elements that include those that appear in more than $1/k$ fraction of the stream.

Hint:

In the majority algorithm, we make bundles of size two. Here, think about bundles of size k.

3 Morris's Approximate Counting Algorithm

In Morris's approximate counting algorithm, which we discussed in the class, prove that

$$\mathbb{E}[2^{2X_m}] = \frac{3}{2}m^2 + \frac{3}{2}m + 1.$$

4 Pairwise Independent Hashing

Prove that given a prime number p, the random hash function $h_{a,b}(x) : \mathbb{F}_p \to \mathbb{F}_p$ defined as $h_{a,b}(x) = ax + b \ mod \ p$, where a and b are random numbers in $\{0, 1, \ldots, p-1\}$, is a pairwise independent hash function. That means, for every $x, y \in \{0, 1, \ldots, p-1\}$ where $x \neq y$, and every $i, j \in \{0, 1, \ldots, p-1\}$, we have $\Pr_{a,b}[h_{a,b}(x) = i \ and \ h_{a,b}(y) = j] = 1/p^2$.

Hint:

Fixing x, y, i, j gives us a system of linear equations on $\mathbb{F}_p$ with two equations and two variables (a and b). Prove that this system always has a unique solution.