Homework 8

Instructor: Brendan Juba

Reminder: You may work in groups and use outside sources. But, you must write up solutions in your own words and properly reference your sources for each problem. This includes listing your collaborators and properly citing any sources you use. Solutions to each problem must be electronically typeset and submitted online via Gradescope. Instructions appear in the E-Homework Guide: http://www.cse.wustl.edu/~bjuba/cse347/s20/ehomework/

1. You are monitoring an aggregation of several Twitter feeds, which is composed of a sequence of $k$-character messages (“tweets”). You are trying to count the number of retweets – tweets that are repeats of some tweet that appeared previously over the course of some given, large number $N$ of tweets ($N$ is greater than the number of distinct characters that may appear in a tweet; we may thus represent these characters as numbers in the range 0–$N$). Give a randomized algorithm that with probability $1 - \delta$ returns a count of the number of retweets among the first $N$ tweets. Your algorithm should only use $O(k)$ arithmetic operations per tweet and only use a polynomial amount of memory (in $1/\delta$, $k$, and $N$).

   You may assume that you are given a subroutine that given $b$, returns a $b$-bit prime number in $O(1)$ time; that you can choose random $b$-bit integers in $O(1)$ time; and that any amount of memory can be allocated initialized to 0 in $O(1)$ time. As usual, you must prove that your algorithm meets the stated guarantees.

2. Kleinberg & Tardos Chapter 13, question 4

3. Kleinberg & Tardos Chapter 8, question 32