CS 537: *Randomness in Computing* — Prof. Sofya Raskhodnikova
Boston University — November 12, 2024

Homework 10 – Due Thursday, November 14, 2022.

Page limit You can submit **at most** 2 pages per problem, even if the problem has multiple parts. If you submit a longer solution for some problem, only the first two pages will be graded.

Reminder Collaboration is permitted, but you must write the solutions *by yourself without assistance*, and be ready to explain them orally to the instructor if asked. You must also identify your collaborators and whether you gave help, received help, or worked something out together. Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden.

Exercises Please practice on exercises in Chapter 5 of Mitzenmacher-Upfal.

Problems

1. (**Hashing with open addressing**) Exercise 5.22 (all parts).

   Clarifications:

   (a) Note that k in part (a) is **not** the same k as in $h(k, j)$ in the preceding paragraph. The authors should have used a different letter here.

   (b) The logarithm in part (b) is base 2.

2. (**Constructions of hash functions**)

   In class (see also Section 15.3.1 of MU), we showed that when p is a prime and $p \geq n$, the family $\mathcal{H} = \{h_{a,b} \mid a \in [p-1], b \in \{0\} \cup [p-1]\}$ is universal, where $h_{a,b}(x) = ((ax + b) \bmod p) \bmod n$.

   Consider the hash functions $h_a(x) = (ax \bmod p) \bmod n$ and the family $\mathcal{H}' = \{h_a \mid a \in [p-1]\}$.

   - **(Don't hand in)** Consider $p = 7$ and make a 7×6 table with rows indexed by x's and columns indexed by a's, showing the values of $ax \bmod p$ for all $x \in \{0\} \cup [p-1]$ and $a \in [p-1]$. What do you observe about each row and each column?

   (a) Let $n = 4$. Change the table from the previous part, so that each cell shows $h_a(x)$.

   (b) Is $\mathcal{H}'$ universal? Explain.

   Now let's consider the general case of $p \geq n$.

   (c) Prove that $\mathcal{H}'$ is almost universal in the following sense: for all $x, y \in \{0\} \cup [p-1]$, if h is chosen uniformly at random from $\mathcal{H}'$ then

   $$\Pr[h(x) = h(y)] \leq \frac{2}{n}.$$

   Hint 1: What condition has to hold for x and y to collide for some value of a?

   Hint 2: Prove that, for fixed x and y, the value of $(ay \bmod p) - (ax \bmod p)$ uniquely determines a.

3. (**Using hashing for comparing multisets**) Exercise 15.11. Define n to be the sum of the sizes of the two multisets.

   For part (a), the book asks to extend this algorithm to a Monte Carlo algorithm. It is confusing, as the algorithm is already Monte Carlo. Instead, explain how to get an algorithm with error probability δ. (In particular, state the value of c you choose.) Try to get the best running time in terms of n and δ.

4. (**MaxSat**) Exercise 6.1. You may assume that m is divisible by 2^k.