CSE 521: Design and Analysis of Algorithms **Fall 2025**

Problem Set 2

Deadline: Oct 15 (at 11:59 PM) in gradescope

Instructions

- You should think about each problem by yourself for at least an hour before choosing to collaborate with others.
- You are allowed to collaborate with fellow students taking the class in solving the problems. But you **must** write your solution on your own.
- You are not allowed to search for answers or hints on the web. You are encouraged to contact the instructor or the TAs for a possible hint.
- You cannot collaborate on Extra credit problems
- Solutions typeset in LATEX are preferred.
- Feel free to use the Discussion Board or email the instructor or the TAs if you have any questions or would like any clarifications about the problems.
- Please upload your solutions to Gradescope.

In solving these assignments and any future assignment, feel free to use these approximations:

$$1-x \approx e^{-x}, \qquad \sqrt{1-x} \approx 1-x/2, \qquad n! \approx (n/e)^n, \qquad \left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k$$

1) (a) For a prime p we generate a pairwise independent hash function by choosing a, b independently from the interval $\{0, \ldots, p-1\}$ and using $ax+b$ as a random number (see lecture 4). Suppose we generate t pseudo random numbers this way, $r_1, \ldots, r_t$ where $r_i = ai + b \pmod{p}$. We want to say this set is far from being mutually independent. Consider the set $S = \{p/2, \ldots, p-1\}$ which has half of all elements. Prove that with probability at least $\Omega(1/t)$ none of the pseudo-random-numbers are in S. Note that if we had mutual independence this probability would have been $1/2^t$.

   (b) Now, assume we have a three-wise independent hash-function $ax^2 + bx + c \bmod p$, for a, b, c chosen independently in the range $[p]$ and suppose we generate again $r_1, \ldots, r_t$ where $r_i = ai^2 + bi + c \bmod p$. What lower bound can you prove on the probability that none of the numbers are in S?

2) Consider the following process for executing n jobs on n processors. In each round, every (remaining) job picks a processor uniformly and independently at random. The jobs that have no contention on the processors they picked get executed, and all other jobs *back off* and then try again. Jobs only take one round of time to execute, so in every round all the processors are available.

   For example, suppose we want to run 3 jobs on 3 processors. Suppose in round 1, jobs 1 and 2 choose the first processor and job 3 chooses the second processor. Then job 3 will be executed and jobs 1 and 2 back off. Suppose in round 2, job 1 chooses the third processor and job 2 chooses the first processor. Then both of them are executed and the process ends in 2 rounds.

   In this problem we almost show that the number of rounds until all jobs are finished is $O(\log \log n)$ with high probability.

a) Suppose less than $\sqrt{n}$ jobs are left at the beginning of some round. Show that for some constant $c > 0$, with probability at least c no job remains after this round.

b) In the first round we have n jobs. Show that the expected number of processors that are picked by no jobs is $n(1 - 1/n)^n$.

c) Suppose there are r jobs left at the beginning of some round. What is the expected number of processors that are matched to exactly one job? What is the expected number of jobs remaining to be completed after that round?

d) Suppose in each round the number of jobs completed is exactly equal to its expectation. Show that (under this false assumption) the number of rounds until all jobs are finished is $O(\log \log n)$.

e) In this part we almost justify the false assumption. Suppose there are r jobs left at the beginning of some round. Let E be the expected number of processors that are matched to exactly one job. Show that for any $k > 1$, the number of processors with exactly one matched job is in the interval $[E - k\sqrt{r}, E + k\sqrt{r}]$ with probability at least $1 - \exp(-\Omega(k^2))$. You can use the McDiarmid's inequality to prove the claim.

Theorem 2.1 (McDiarmid's inequality)**.** *Let $X_1, \ldots, X_n \in \mathcal{X}$ be independent random variables. Let $f : \mathcal{X}^n \to \mathbb{R}$. If for all $1 \leq i \leq n$ and for all $x_1, \ldots, x_n$ and $\tilde{x}_i$,*

$$|f(x_1, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, \tilde{x}_i, x_{i+1}, \ldots, x_n)| \leq c_i,$$

then

$$\mathbb{P}\left[|f(X_1, \ldots, X_n) - \mathbb{E}\left[f\right]| \geq \epsilon\right] \leq 2 \exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^n c_i^2}\right)$$