Prof. Juraj Hromkovič
Juraj.Hromkovic@inf.ethz.ch
Richard Královič
Richard.Kralovic@inf.ethz.ch

Design of Randomized Algorithms

Exercises

Exercise 1

Def. 1 Definition of protocol R_2.

Protocol R_2 is a modified randomized protocol for equity that chooses two primes, p and q, at random, and accepts an input (x, y) if and only if

$$(x \bmod p = y \bmod p) \wedge (x \bmod q = y \bmod q).$$

Def. 2 Definition of the random variable Y.

$$Y = \begin{cases} 1 & \text{if } p \text{ or } q \text{ is good for } (x,y) \\ 0 & \text{if both } p \text{ and } q \text{ are bad for } (x,y) \end{cases}$$

Consider the probabilistic experiment given by the protocol R_2. Define two indicator variables X_1 and X_2, where X_1=1 if and only if the first prime p is good for (x, y) and X_2=1 if and only if the second prime q is good for (x, y). Prove that X_1 and X_2 are independent. Can X_1 and X_2 be used to define the random variable Y as used above?

Exercise 2

The randomized protocol (called d-R) for equity that chooses witnesses among the primes from range 1 to n^d has probability of error at most $\frac{d \ln n}{n^{d-1}}$ and communication complexity at most $2d\lceil \log_2 n \rceil$ (for n large enough).

Compare the amplification method of executing independent runs with the method of increasing the size of the set of witness candidates (the set of fingerprinting methods). How many communication bits are needed by each of these methods in order to get an error probability tending to 0 with growing n? By what size is the error probability reducible by these methods when an upper bound $c(n)$ on the number of communication bits is given?

Exercise 3

Assume that the assumptions of Lemma 3.2.2 are satisfied. Estimate the expected number of collisions in all slots of T.

Lemma 3.2.2. *Let $U = \mathbb{N}$ be the universe, and let $T = \{0, 1, \ldots, m-1\}$, $m \in \mathbb{N} - \{0, 1\}$. Let n be a positive integer, and let $h : U \to T$ be a hash function that satisfies (3.1), i.e. $Prob(h(x) = i) = \frac{1}{m}$ for all $0 \leq i \leq m-1$. Then, for every slot l of T,*

(i) *the expected number of elements of a random $S \in \mathcal{P}_n(U)$ assigned to the slot l (i.e., the number of elements x of S with $h(x) = l$) is smaller than*

$$\frac{n}{m} + 1$$

(ii) *and, if $n = m$, then*

$$Prob(\text{more than one key from a random } S \in \mathcal{P}_n(U) \text{ is in the } l\text{-th slot}) < \frac{1}{2}.$$

Exercise 4

The example of a universal class of hash functions in the book[1] is based on the fact that $T = \{0, 1, \ldots, m-1\}$ for a prime m. Consider $m = p^a$ for a prime p and a positive integer $a \geq 2$. For such an m, does there exist a universal class of hash functions?

Exercise 5

Consider the problem of finding k-th smallest element of a given set $A = \{a_1, \ldots, a_n\}$. The input of the problem is a pair (A, k). One possible solution is to sort the set A, what yields solution with complexity $\Theta(n \log n)$. Another possible solution is the following randomized algorithm:

Algorithm RS$(\mathbf{A}, \mathbf{k})$

Input: $A = \{a_1, \ldots, a_n\}$ for integer $n \in \mathbb{N}^{>0}$ and integer k such that $1 \leq k \leq n$.

Step 1: If $n = 1$, output "a_1", otherwise choose $i \in \{1, \ldots, n\}$ randomly.

Step 2: Compute $A_< = \{b \in A \mid b < a_i\}$ and $A_> = \{c \in A \mid c > a_i\}$.

Step 3: If $|A_<| > k$, then call $RS(A_<, k)$; if $|A_<| = k - 1$, then output "a_1"; otherwise call $RS(A_>, k - |A_<| - 1)$.

Obviously $RS(A, k)$ gives always correct answer. Prove that the expected time complexity is linear: Exp-Time$_{RS}(A, k) \in O(n)$

Exercise 6

Consider the following modification of Algorithm RSAM:

Input: Formula $\Phi = F_1 \wedge F_2 \wedge \ldots \wedge F_m$ in CNF over variables $\{x_1, x_2, \ldots, x_n\}$.

Step 1: Choose random assignment of truth values $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ for variables $x_1, x_2, \ldots, x_n$.

Step 2: Compute number $r(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of clauses satisfied by assignment $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. If $r(\alpha_1, \alpha_2, \ldots, \alpha_n) \geq \frac{m}{2}$, then output $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, otherwise repeat Step 1.

If this algorithm halts, it computes an assignment that satisfies at least one half of the clauses. What is the expected number of repeats of Step 1?

Exercise 7

We choose t random assignment of truth values to the variables of Formula Φ in CNF. What is the probability that at least one of these assignments satisfies at least one half of the clauses of Φ?

[1] Universe is defined as $(r+1)$-dimensional vector space over field $\mathbb{Z}_m$, the hash function $h_\alpha(x)$ is defined as $\left(\sum_{i=0}^{r} \alpha_i x_i\right) \bmod m$.

Exercise 8

Modify the 2MC Algorithm A_t to a randomized algorithm A'_t in such a way that A'_t takes the most frequent result as the output What is the error probability of this modified algorithm A'_t?

Exercise 9

Apply the amplification method in order to enable a randomized test of $x \in U$ for larger sets U. Estimate the maximal possible cardinality of U for which the achieved error probability still approaches 0 with growing n when the communication complexity is

(i) in $O(\log n \cdot \log\log n)$,

(ii) in $O((\log n)^d)$ for some constant $d \in \mathbb{N}$,

(iii) polylogarithmic

Exercise 10

Consider the following communication problem: Given two sets U, V, decide if U and V are disjoint.

Let $U, V \subseteq \{0,1\}^n$ such that $|U| \in O(n^3)$ and $|V| \in O(n^2)$. How many times do we need to repeat protocol PDisj to make error probability converging to 0? What communication complexity do we obtain? What happens in the case when $U, V \in \Theta(2^{\sqrt{n}})$?

Exercise 11

Let $U = \mathbb{N}$ be the universe, and let $T = \{0, 1, \ldots, m-1\}$, $m \in \mathbb{N} - \{0,1\}$. Let $h : U \to T$ be a hash function such that $Prob(h(x) = i) = \frac{1}{m}$ for all $i \in T$.

Let $n \in \mathbb{N} - \{0,1\}$ and let S be a randomly chosen subset of U containing n elements. What is the expected number of collisions in all m slots of T?

Exercise 12

Let $U = \mathbb{N}$ be the universe, and let $T = \{0, 1, \ldots, m-1\}$, $m \in \mathbb{N} - \{0,1\}$. Consider the set of all hash functions $h : U \to T$ such that for each $i \in T$ it holds that

$$Prob(h(x) = i) = \frac{1}{m}.$$

Is M a universal set of hash functions?

Exercise 13

Let $U = \{0, 1, 2, \ldots, p-1\}$ be a universe for some prime number p. Let $T = \{0, 1, \ldots, m-1\}$. For arbitrary positive integers $a, b \in U$, we define the linear hash function $h_{a,b}: U \to T$ as follows:

$$h_{a,b}(x) = ((ax + b) \bmod p) \bmod m.$$

Let $H^p_{\text{lin}} = \{h_{a,b} \mid a \in \{1, 2, \ldots, p-1\} \text{ and } b \in \{0, 1, \ldots, p-1\}\}$. In the lecture we have proved that H^p_{lin} is a universal set of hash functions $U \to V$.

Let l be an arbitrary positive integer and p' be arbitrary prime number $p' > l$. Consider the set $U' = \{0, 1, \ldots, l\}$. Is $H^{p'}_{\text{lin}}$ a universal set of hash functions $U' \to T$?