Assignment 5

Exercise 5.1. (8)  
_Making random sources uniformly distributed_

Our task is to build a random source that outputs the bits 0 and 1 with \( \text{prob}(0) = \text{prob}(1) = \frac{1}{2} \). We have access to another random source \( S \) that outputs \( a \) or \( b \) with independent probabilities \( \text{prob}(a) \) and \( \text{prob}(b) = 1 - \text{prob}(a) \) that are unknown to us.

State an algorithm that does the job and that does not consume more than an expected number of \( (\text{prob}(a) \cdot \text{prob}(b))^{-1} \) symbols of \( S \) between two output bits. Prove its correctness.

Exercise 5.2. (8)  
_Fingerprinting_

Two processors \( A, B \) with inputs \( a \in \{0, 1\}^n \) (for \( A \)) and \( b \in \{0, 1\}^n \) (for \( B \)) want to decide whether \( a = b \). \( A \) does not know \( B \)'s input and vice versa.

\( A \) can send a message \( m(a) \in \{0, 1\}^* \) which \( B \) can use to decide \( a = b \). The communication and computation rules are called a _protocol_.

- Show that every deterministic protocol must satisfy \( |m(a)| \geq n \).
- State a randomized protocol that uses only \( O(\log_2 n) \) Bits. The protocol should always accept if \( a = b \) and accept with probability at most \( \frac{1}{n} \) otherwise. Prove its correctness.

Exercise 5.3. (8)  
_Continuous uniform samples_

A source provides a stream of items \( x_1, x_2, \ldots \). At each step \( n \) we want to save a random sample \( S \subseteq \{(x_i, i)|1 \leq i \leq n\} \) of size \( k \), i.e. \( S \) should be a uniformly chosen sample from all \( \binom{n}{k} \) possible samples consisting of seen items. So at each step \( n \geq k \) we must decide whether to add the next item to \( S \) or not. If so we must also decide which of the current items to remove from \( S \).

State an algorithm for the problem. Prove its correctness.