

This homework covers Lectures 3 and 4. All arithmetic is modulo 2.

1. Consider the following two samplers:

- X takes in 5 pure random bits $r_0 r_1 r_2 r_3 r_4$ and outputs the 3 bits

$$x_1 = r_0, x_2 = r_1 r_2, x_3 = r_3 r_4.$$

- Y takes in 5 pure random bits $r_0 r_1 r_2 r_3 r_4$ and outputs the 6 bits $r_0 + r_i r_j$, $1 \leq i < j \leq 4$, namely

$$y_1 = r_0 + r_1 r_2, y_2 = r_0 + r_3 r_4, y_3 = r_0 + r_1 r_3, y_4 = r_0 + r_2 r_4, y_5 = r_0 + r_1 r_4, y_6 = r_0 + r_2 r_3.$$

- Calculate the probabilities of the events $x_i = 1$ for all i in sampler X .
- Show that the test "Does x have Hamming weight at most 1?" distinguishes the output of X from pure random bits. What is its advantage?
- (**Optional**) Is there a distinguisher with higher advantage than the one from part (b)?
- Calculate the probabilities of the events $y_i = 1$ for all i in sampler Y .
- As Y 's seed is shorter than its output, its output must be distinguishable from 6 pure random bits. Describe a distinguisher that does the job.

2. This question is about indistinguishability of single versus multiple samples.

- Let S be the sampler that takes in two bits of seed r_1, r_2 and outputs $(r_1, r_2, r_1 + r_2)$. What is the advantage of the best distinguisher that tells the output of S from truly random bits?
- Let S^2 be the sampler that runs S twice on independent seeds: $S^2(r_1, r_2, r'_1, r'_2) = (S(r_1, r_2), S(r'_1, r'_2))$. Here S^2 produces six bits of output. What is the advantage of the best distinguisher for S^2 ?
- Now S is arbitrary. Assume distinguisher D tells S from pure random bits, and D accepts all outputs of S . How should you use D to distinguish S^2 from random? How does the advantage change?
- (**Optional**) Now assume distinguisher D' tells S^2 from pure random bits. How can you use D' to distinguish S from random? (**Hint:** Don't aim for maximum advantage.)

3. R is a pure random function that maps 4 input bits into one output bit. F is the keyed function

Setup stage: Sample and store a random 5-bit key $k = k_0 k_1 k_2 k_3 k_4$.

Evaluation stage: On input $x = x_1 x_2 x_3 x_4$, output $k_0 + k_1 x_1 + k_2 x_2 + k_3 x_3 + k_4 x_4$.

- What are the probabilities of the events $R(x) = 1$ and $R(x) + R(y) = 0$ for distinct inputs $x \neq y$?
- What is the probability that $F(x) = 1$ for a given input x ?
- What is the probability that $F(x) + F(y) = 0$ for a given distinct pair $x \neq y$?
- (**Optional**) If Eve makes only two queries, can she distinguish R from F ?
- What is the probability that $R(x) + R(y) + R(z) + R(w) = 0$ for distinct inputs x, y, z, w ?
- Find inputs for which the probability that $F(x) + F(y) + F(z) + F(w) = 0$ is different from that for R .

4. Assume F is a pseudorandom function with equal input and output length. The secret key to F is known to Alice and Bob only. Diagnose the problem with each of the following variants of the PRF pad encryption:
- (a) m encrypts to $F(r) + m$ only, where r is chosen at random at encryption time.
 - (b) m encrypts to $(r, F(r) + m)$, where r is chosen once at key generation and reused for every encryption.
 - (c) m encrypts to $(r, F(r) + m, F(r + m))$.