Homework 1

January 27, 2004

This problem set is due in class on **Monday, February 2.**

1 Computational indistinguishability

We define a distance measure on probability distributions as follows:

$$d_R(D, D') = \max_A |\Pr_{x \leftarrow D}[A(x) = 1] - \Pr_{x' \leftarrow D'}[A(x') = 1]|$$

where the maximum is taken over all algorithms running with resources R. Here R denotes a set of resources, which is used to restrict the adversary. For instance, R might list a bound on number of steps of computation, on queries to the oracle, on the amount of memory used, or something else. If algorithm A runs with resources R and A' with R', then let $R + R'$ denote the resources used by running first A then A'. If you like, you can assume that $R + R' = R' + R$.

Note that D, D' are (t, e)-indistinguishable if and only if $d_t(D, D') \leq e$, so we can see that this distance measure is of fundamental interest in cryptography. This problem will ask you to study basic properties of this distance measure.

(a) [*Data processing inequality*] Let L be any (possibly randomized) algorithm running with resources R', and let $L(D)$ denote the distribution obtained by picking x from D and then outputting $L(x)$. Prove: $d_R(L(D), L(D')) \leq d_{R'+R}(D, D')$.

(b) [*Triangle inequality*] Let D' be any distribution whatsoever. Prove: $d_R(D, D'') \leq d_R(D, D') + d_R(D', D'')$.

(c) [*Multiple samples*] If D_1, D_2 are distributions, let $D_1 \times D_2$ denote the distribution on pairs whose first component is chosen according to D_1 and second component according to D_2 (independently), i.e., $\Pr_{D_1 \times D_2}[(x_1, x_2)] = \Pr_{D_1}[x_1] \times \Pr_{D_2}[x_2]$.

Prove: $d_R(D \times D, D' \times D') \leq 2 d_R(D, D')$.

(d) [*Information-theoretic security and the variation distance*] Recall that the variation distance is defined by $V(D_1, D_2) = \frac{1}{2} \sum_x |\Pr_{D_1}[x] - \Pr_{D_2}[x]|$. When used as a resource bound, let ∞ denote that there are no restrictions on the adversary (e.g., the algorithm can have unbounded running time).

Prove: $d_\infty(D, D') = V(D, D')$.

(e) [*Information-theoretic security*] Prove: $d_R(D, D') \leq d_\infty(D, D')$.

(f) [*Conditioning*] If E is an event, let $D|E$ represent the distribution of the output of D conditioned on E occurring, i.e., $\Pr_{D|E}[x] = \Pr_D[x|E] = \Pr_D[x \text{ and } E] / \Pr[E]$. Let $\overline{E}$ denote the complementary event to E, i.e., that E does not occur.

Prove: $d_R(D, D') \leq d_R(D|E, D'|E) + \Pr[\overline{E}]$.

(g) [*A XOR lemma*] If D_1, D_2 are distributions on k-bit strings, let $D_1 \oplus D_2$ denote the distribution on $x_1 \oplus x_2$ when x_1 is chosen from D_1 and x_2 is chosen from D_2 (independently). Here $\oplus$ denotes the xor operator. In other words, $\Pr_{D_1 \times D_2}[y] = \sum \Pr_{D_1}[x_1] \times \Pr_{D_2}[x_2]$, where the sum is taken over all pairs (x_1, x_2) such that $x_1 \oplus x_2 = y$. Let U denote the uniform distribution on k-bit strings.

Prove: $d_R(D \oplus D', U) \leq 2 d_\infty(D, U) d_\infty(D', U)$.