Foundations of Cryptography

Oded Goldreich and Shafi Goldwasser

Comments on Exercise 3

Comment A (question 1.a)

Given a zero-knowledge interactive proof system (P, V) for 3-COLORABILITY, you define a new proof system (P', V') for L using the NP reduction f. Specifically, $V'(x) = V(f(x))$. The following proof of soundness is incorrect:

$$\forall x \in L \text{ and p.t.m. } B: \ \Pr\left[\left\langle B, V'\right\rangle(x) = 1\right] = \Pr\left[\left\langle B, V\right\rangle(f(x)) = 1\right] = \cdots$$

The inequality generally does not hold, because the input to B is different. There are several ways to fix this, one of which is as follows: we can assume without loss of generality that B ignores its input, since if it doesn't then we can replace it by $\bar{B}$ that is identical to B except it has the input x hard-wired.

Comment B (question 1.b.)

Given a zero-knowledge interactive proof system (P, V) for 3-COLORABILITY, you define a new proof system (P', V') for L. You now need to show that the latter is also zero-knowledge. You *cannot* use an argument like "let S be the simulator for the prover P, ..." — the *existence* of simulators is indeed a property of P, but each cheating verifier V^* has a *different* simulator S. This is due to the order of quantifiers in the definition of zero knowledge, and indeed cannot be otherwise because the distribution of interactions $\langle P, V^* \rangle(x)$ clearly depends on V^*. Thus, to complete the proof you need to define a *specific* V^* (as a function of a cheating V'^* that interacts with P') and use the existence of a simulator for $\langle P, V^* \rangle$.

Comment C (question 2.a.)

Given a public key cryptosystem (G, E, D) that has indistinguishability security, you need to define a one-way function. A common answer was to define $f(m, r_1, r_2) = E_e^{r_1}(m)$ for $(e, d) = G^{r_2}(1^{|m|})$, where the superscript denotes fixing the random coins. It was then argued that if there exists an inverter I such that $\Pr_{m,r_1,r_2}\left[I(f(m, r_1, r_2)) \in f^{-1}f(m, r_1, r_2)\right]$ is not negligible, then using I we can define a machine D that distinguishes encryptions, hence contradicting the security of the cryptosystem. This is incorrect, for the following reason. I needs to find *some* triplet (m', r_1', r_2') such that $f(m', r_1', r_2') = f(m, r_1, r_2)$. Since we allow $r_1' \neq r_1$ (i.e., different encryption keys), I may (nearly) always return a triplet with $m' \neq m$, and thus D will (nearly) always fail. In this case, there is no contradiction to indistinguishability securiye.

To fix this, you need to force f to return a "relevant" answer by adding the public key to its output: $f(m, r_1, r_2) = (E_e^{r_1}(m), e)$ for $(e, d) = G^{r_2}(1^{|m|})$, or equivalently $f(m, r_1, r_2) = (E_e^{r_1}(m), r_1)$.

Comment D (question 2.b.)

Given a public key cryptosystem (G, E, D) that has indistinguishability security, you need to define a collection of trapdoor predicate. All of the submitted proofs are incorrect in two ways, as the following typical variant illustrates (several variants were suggested, but they all fail in the same ways). Consider the following trapdoor predicate (I use the notation of Definition 2.59 on page 3.9 in the lecture notes of Goldwasser and Bellare, except D_i is replaced by $\mathcal{D}_e$):

- $I = \{e \mid \exists n, r : G^r(1^n) = (e, d)\}$ where the superscript denotes fixing the random coins
- $\mathcal{D}_e = E_e(m_0) \cup E_e(m_1)$ where $m_0 = 0^k$, $m_1 = 1^k$, $k = |e|$
- $B_e(c) = \begin{cases} 0 & \text{if } D_d(c) = m_0 \\ 1 & \text{otherwise} \end{cases}$ where d is the decryption key corresponding to e

Now, suppose there exists a PPT A such that the following probability is not negligible:

$$\Pr\left[z = v \mid e \in_R I \cap \{0,1\}^k,\ v \in_R \{0,1\},\ x \in_R \mathcal{D}_e^v,\ z \leftarrow A(e, x)\right]$$

In an obvious manner, $A(e, x)$ can help in distinguishing encryptions of m_0 (i.e., $x \in \mathcal{D}_e^0$) from encryptions of m_1 (i.e., $x \in \mathcal{D}_e^1$). Does this contradict the security of (G, E, D)? Not necessarily, because if we compare the above to the definition of indistinguishability security we see that some probability distributions are different between the two: the distribution of keys and the distribution of ciphertexts. There, they are defined by the algorithms G and E respectively. Here, they are "flattened" into the uniform distribution over the support sets I and D respectively. Details follow.

First, we consider the distibution of keys. Fix k, and suppose that there exists a ciphertext c and a set of keys S such that for all keys in S, c decrypts into m_0 and for all other keys, c decrypts into m_1. Is it possible that S is much larger than $\bar{S}$, e.g., $|S| \approx 10\left|\bar{S}\right|$? There's no immediate reason why not, because each specific key in S may have a lower probability of being generated by G than each specific key in $\bar{S}$. If this is indeed the case then $I(e, x)$

can always return 0, and be right with probability 10/11 (but see the next point). Thus the argument fails.

A similar problem arises with respect to the probability of the ciphertexts. Both m_0 and m_1 must have an exponential number of possible encryptions (why?), but it's quite possible, for instance, that m_0 has many more possible encryptions than m_1: $\forall e : |E_e(m_0)| \approx 10\,|E_e(m_1)|$. Of course, the average probability of a *specific* encryption of $m_0$being generated by E would be about $10 times smaller than for m_1$, but this probability distribution is “flattened” in the defininition of trapdoor predicates. Thus in this case, $I(e, x)$ can again always return 0 and be right with probability 10/11, yet there is no immediate contradiction to indistinguishability security.

There does not appear to be a simple way to fix these other than changing the definition of collections of trapdoor predicates, or assuming that the distribution of keys generated by G and the distribution of ciphertexts generated by E are uniform.

The moral of this story is that when arguing about probability, the explicit specification of the distribution over which the probability is taken is not always a mere technicality.

Comment E (question 3.c)

To prove the security of Phila's cryptosystem we use a hybrid argument, i.e., we present a sequence of ensembles such that if some pair of adjacent ensembles is not computational indistinguishabile then we obtain a contradiction. Here is a sketch of a correct argument. Let $\stackrel{D}{\equiv}$ denote computational indistinguishability. For any $m \in \{0,1\}^k$:

$$
\begin{aligned}
& \{e, E_e(r), Q(r) \oplus m\}_{(e,d) \leftarrow G(1^{k/2}), r \in \{0,1\}^{k/2}, \text{coins}} \\
\stackrel{D}{\equiv} \quad & \text{(by } (G, E, D) \text{ having indistinguishability security)} \\
& \left\{e, E_e(0^{k/2}), Q(r) \oplus m\right\}_{(e,d) \leftarrow G(1^{k/2}), r \in \{0,1\}^{k/2}, \text{coins}} \\
\stackrel{D}{\equiv} \quad & \text{(by } Q \text{ being a PRG; it's essential that } r \text{ does not appear more than once)} \\
& \left\{e, E_e(0^{k/2}), U_k \oplus m\right\}_{(e,d) \leftarrow G(1^{k/2}), \text{coins}} \\
\equiv \quad & \\
& \left\{e, E_e(0^{k/2}), U_k\right\}_{(e,d) \leftarrow G(1^{k/2}), \text{coins}}
\end{aligned}
$$

By transitivity it follows that $\forall m_0, m_1 \in \{0,1\}^n : \{e, E'_e(m_0)\}_{e,\text{coins}} \stackrel{D}{\equiv} \{e, E'_e(m_1)\}_{e,\text{coins}}$, which is equivalent to indistinguishability security.

In the first two transitions, we use the preservation of computational indistinguishability under efficient transformations (property from Exercise 2 question 1.a). The order of the hybrids matters: we cannot use the PRG property of $Q(r)$ when r appears elsewhere, because then there is no suitable efficient transformation.

Comment F (question 3.c)

To prove indistinguishability security of (G', E', D'), you need to prove computational indistinguishability of ensembles that include not only the ciphertext, but also the encryption key e. Consult the definition of indistinguishability security to see why. Note that the addition of e makes the claim stronger, and that the stronger claim generally does not follow from the weaker one using preservation of computational indistinguishability under efficient transformations. Comment E above sketches a correct proof.

Comment G (question 5.b)

For message tuples $\bar{m}_0, \bar{m}_1$ and a random permutation g, the distributions of encryptions of $\bar{m}_0$ is *not* necessarily identical to the distribution of encryptions of $\bar{m}_1$. Consider, for instance, the case where $m_{0,1} = m_{0,2}$ and $m_{1,1} \neq m_{1,2}$. The first two ciphertexts will sometimes be equal for $\bar{m}_0$, but can never be equal for $\bar{m}_1$. You need to argue that this happens with negligible probability.

Note that a different way to answer this question is by invoking question 5.a. To do so, we need to reduce multiple-messages indistinguishability to a chosen-plaintext attack, which can be done using a hybrid argument similar to the one seen in class. Note that the availability of an encryption oracle is crucial to the hybrid argument.

Questions may be directed to Eran Tromer `<tromer@wisdom.weizmann.ac.il>`.