

Problem Set 2 — Solutions

1. We first show that if a scheme is not perfectly secret, then it is not perfectly indistinguishable. Assume the scheme is not perfectly secret. Then, by Lemma 2.3 there exist messages $m_0, m_1 \in \mathcal{M}$ and a ciphertext $\bar{c} \in \mathcal{C}$ such that

$$\Pr[C = \bar{c} \mid M = m_0] \neq \Pr[C = \bar{c} \mid M = m_1].$$

Consider the following adversary \mathcal{A} :

- (a) Output m_0, m_1 (where these are the messages guaranteed to exist as above).
- (b) When given a ciphertext c (that is either an encryption of m_0 or an encryption of m_1): if $c = \bar{c}$ then output $b = 0$; else output $b' = 1$.

We have

$$\begin{aligned} \Pr[\text{PrivK}_{\mathcal{A}, \Pi}^{\text{eav}} = 1] &= \Pr[b' = b] \\ &= \Pr[b' = b \wedge b = 0] + \Pr[b' = b \wedge b = 1] \\ &= \frac{1}{2} \cdot \Pr[b' = 0 \mid b = 0] + \frac{1}{2} \cdot \Pr[b' = 1 \mid b = 1], \end{aligned}$$

just using the rules of probability. Now, notice that \mathcal{A} outputs $b' = 0$ only if $c = \bar{c}$. Furthermore, $b = 0$ means that m_0 was encrypted in the experiment, and likewise $b = 1$ means that m_1 was encrypted. So:

$$\begin{aligned} \Pr[\text{PrivK}_{\mathcal{A}, \Pi}^{\text{eav}} = 1] &= \frac{1}{2} \cdot \Pr[b' = 0 \mid b = 0] + \frac{1}{2} \cdot \Pr[b' = 1 \mid b = 1] \\ &= \frac{1}{2} \cdot \Pr[c = \bar{c} \mid m_0 \text{ encrypted}] + \frac{1}{2} \cdot \Pr[c \neq \bar{c} \mid m_1 \text{ encrypted}] \\ &= \frac{1}{2} \cdot \Pr[C = \bar{c} \mid M = m_0] + \frac{1}{2} \cdot (1 - \Pr[C = \bar{c} \mid M = m_1]) \\ &= \frac{1}{2} + \frac{1}{2} \cdot (\Pr[C = \bar{c} \mid M = m_0] - \Pr[C = \bar{c} \mid M = m_1]). \end{aligned}$$

Since $\Pr[C = \bar{c} \mid M = m_0] \neq \Pr[C = \bar{c} \mid M = m_1]$, the second term above cannot be 0 and so $\Pr[\text{PrivK}_{\mathcal{A}, \Pi}^{\text{eav}} = 1] \neq \frac{1}{2}$. This shows that Π is not perfectly indistinguishable.

Next, we show that if a scheme is perfectly secret then it is perfectly indistinguishable. Assume the scheme is perfectly secret and fix an adversary \mathcal{A} . We do not know anything about \mathcal{A} , but we will still show that $\Pr[\text{PrivK}_{\mathcal{A}, \Pi}^{\text{eav}} = 1] = \frac{1}{2}$. We begin by writing:

$$\Pr[\text{PrivK}_{\mathcal{A}, \Pi}^{\text{eav}} = 1] = \frac{1}{2} \cdot \Pr[b' = 0 \mid b = 0] + \frac{1}{2} \cdot \Pr[b' = 1 \mid b = 1]. \quad (1)$$

(This is as above.) Let m_0, m_1 be the two messages output by \mathcal{A} , and let C denote the ciphertext given to \mathcal{A} in the experiment. Looking at the first term (above) by itself, we have:

$$\begin{aligned} \Pr[b' = 0 \mid b = 0] &= \Pr[b' = 0 \mid M = m_0] \\ &= \sum_{c \in \mathcal{C}} \Pr[b' = 0 \wedge C = c \mid M = m_0] \\ &= \sum_{c \in \mathcal{C}} \Pr[C = c \mid M = m_0] \cdot \Pr[b' = 0 \mid C = c \wedge M = m_0]. \end{aligned}$$

Now, the key point is that *the output b' of \mathcal{A} is determined only by C* (because \mathcal{A} is only given the ciphertext, not the message). So,

$$\begin{aligned} \Pr[b' = 0 \mid b = 0] &= \sum_{c \in \mathcal{C}} \Pr[C = c \mid M = m_0] \cdot \Pr[b' = 0 \mid C = c \wedge M = m_0] \\ &= \sum_{c \in \mathcal{C}} \Pr[C = c \mid M = m_0] \cdot \Pr[b' = 0 \mid C = c]. \end{aligned}$$

Similarly, we obtain

$$\Pr[b' = 1 \mid b = 1] = \sum_{c \in \mathcal{C}} \Pr[C = c \mid M = m_1] \cdot \Pr[b' = 1 \mid C = c].$$

Because the scheme is perfectly secret, we know that for any $c \in \mathcal{C}$:

$$\Pr[C = c \mid M = m_0] = \Pr[C = c] = \Pr[C = c \mid M = m_1].$$

Let $\gamma_c \stackrel{\text{def}}{=} \Pr[C = c]$. Plugging the last three equations into Equation (1), we get

$$\begin{aligned} \Pr[\text{PrivK}_{\mathcal{A}, \Pi}^{\text{eav}} = 1] &= \frac{1}{2} \cdot \left(\sum_{c \in \mathcal{C}} \gamma_c \cdot \Pr[b' = 0 \mid C = c] + \sum_{c \in \mathcal{C}} \gamma_c \cdot \Pr[b' = 1 \mid C = c] \right) \\ &= \frac{1}{2} \cdot \sum_{c \in \mathcal{C}} \gamma_c \cdot (\Pr[b' = 0 \mid C = c] + \Pr[b' = 1 \mid C = c]) \\ &= \frac{1}{2} \cdot \sum_{c \in \mathcal{C}} \gamma_c = \frac{1}{2}, \end{aligned}$$

since the only possibilities are $b' = 0$ or $b' = 1$, and $\sum_{c \in \mathcal{C}} \Pr[C = c] = 1$.

2. This is false. This is obvious if you think about it, but to see it formally note that for any perfectly-secure scheme we have

$$\begin{aligned} \Pr[M = m \mid C = c] &= \frac{\Pr[M = m \wedge C = c]}{\Pr[C = c]} = \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\Pr[C = c]} \\ &= \Pr[M = m], \end{aligned}$$

using Lemma 2.2. Similarly, $\Pr[M = m' \mid C = c] = \Pr[M = m']$. Taking any distribution where the *a priori* probabilities of m and m' are different shows that these will not, in general, be equal.

3. (a) The proof is identical to the proof for the one-time pad, with the only difference being that all addition is now occurring modulo 26 rather than modulo 2.
- (b) Let Σ denote the English alphabet. The substitution cipher is certainly perfectly secret when the message space is Σ . But you can do better. The thing to note is that, when encrypting a multi-character message, the only potential problem is a repeated character; otherwise the scheme is perfectly secret. (This is not a formal proof, but one can be given.) So we can take our message space to be

$$\mathcal{M} = \{w \in \Sigma^{26} \mid w \text{ has no repeated characters}\}.$$

Note that $|\mathcal{M}| = 26!$.

Can you find a larger message space for which perfect secrecy still holds? Can you *prove* that this is impossible? (*Hint*: what is the size of the key space?)

- (c) If the Vigenère cipher with a key of length t is used (where t is the length of the messages being encrypted), then it is perfectly secret. As in part (a), a proof is essentially identical to the proof for the one-time pad.

Reconcile this with the attacks that were shown in class: The attacks shown in class all assumed that longer plaintexts were being encrypted. The above only claims perfect secrecy for relatively short plaintexts.

4. (Please read and understand the clarification of this [and the next] question posted on the webpage.) As the hint suggests, take $m \neq m'$ with $\Pr[M = m \wedge M' = m'] \neq 0$, and arbitrary *equal* ciphertexts c, c' . Then

$$\Pr[M = m \wedge M' = m' \mid C = c \wedge C' = c'] = 0$$

because the encryption of two different messages cannot yield the same ciphertext (if it did, then there would be a decryption error in one case or the other). Put otherwise, conditioned on seeing the *same* ciphertext sent twice over the channel, an eavesdropper knows that the corresponding messages must have been the same.

5. Consider the one-time pad for *single-bit* messages. You should be able to convince yourself intuitively that it works. Formally, fix non-equal m, m' and non-equal c, c' (these values m, m', c, c' are all just bits). Then observe that

$$\Pr[C = c \wedge C' = c' \mid M = m \wedge M' = m'] = \Pr[k = m \oplus c] = \frac{1}{2}$$

and

$$\begin{aligned} \Pr[C = c \wedge C' = c'] &= \Pr[C = c \wedge C' = c' \mid M \neq M'] \cdot \Pr[M \neq M'] \\ &= \frac{1}{2} \cdot \Pr[M \neq M'] \end{aligned}$$

using the fact that $C \neq C'$ cannot occur if $M = M'$. Then, just as in the book:

$$\Pr[M = m \wedge M' = m' \mid C = c \wedge C' = c']$$

$$\begin{aligned}
&= \frac{\Pr[M = m \wedge M' = m' \wedge C = c \wedge C' = c']}{\Pr[C = c \wedge C' = c']} \\
&= \frac{\Pr[C = c \wedge C' = c' \mid M = m \wedge M' = m'] \cdot \Pr[M = m \wedge M' = m']}{\Pr[C = c \wedge C' = c']} \\
&= \frac{\Pr[M = m \wedge M' = m']}{\Pr[M \neq M']} = \frac{\Pr[M = m \wedge M' = m' \wedge M \neq M']}{\Pr[M \neq M']}.
\end{aligned}$$

But this is just $\Pr[M = m \wedge M' = m' \mid M \neq M']$.

More generally, for any message space \mathcal{M} if we let the key be a random permutation of \mathcal{M} (with encryption done in the natural way) then we get a scheme satisfying the definition. (This is exactly a substitution cipher over the large ‘alphabet’ \mathcal{M} .) This means the size of the key-space \mathcal{K} satisfies $|\mathcal{K}| = |\mathcal{M}|!$. It *is* possible to do better, but I leave this as a challenge for the reader.

6. As in the proof of Theorem 2.7, we can show that if $|\mathcal{K}| < |\mathcal{M}|$ then there exist messages m_0, m_1 and a ciphertext c such that c is an encryption of m_0 (under some key) but c is *not* an encryption of m_1 (under any key). That is,

$$\Pr[C = c \mid M = m_1] = 0 < \Pr[C = c \mid M = m_0].$$

Then proceeding exactly as in problem 1 gives the desired adversary along with an analysis.

7. Recall the definition of negligible: the function ϵ is negligible if for any polynomial p there exists an integer N such that for all $n > N$ we have $\epsilon(n) < 1/p(n)$.

So we need to prove that for every polynomial p there exists an N as above. Fix an arbitrary polynomial p . Since ϵ_1 is negligible and $2p$ is still polynomial, there exists an N_1 such that for all $n > N_1$ we have $\epsilon_1(n) < 1/2p(n)$. Similarly, there exists an N_2 such that for all $n > N_2$ we have $\epsilon_2(n) < 1/2p(n)$. Take $N = \max\{N_1, N_2\}$. Then for any $n > N$ it holds that

$$\epsilon(n) = \epsilon_1(n) + \epsilon_2(n) < \frac{1}{2p(n)} + \frac{1}{2p(n)} = \frac{1}{p(n)},$$

using the fact that $N \geq N_1, N_2$. This is exactly what we needed to show.