

Solutions to homework 1

1 Substitution cipher [15 points]

Have a look at the substitution cipher in Lecture Notes 1 (section 3.2) and recall the definition of perfect secrecy. Prove that the substitution cipher is perfectly secure for the special case of $\ell = 1$, and that it is *not* perfectly secure if $\ell \geq 2$.

Solution: This cipher meet the perfect secrecy requirement for $\ell = 1$, because for every $m, c \in \{A, \dots, Z\}$ exactly one out of every 26 permutations k between letters maps letter m to letter c , and hence $\text{Prob}_{k \in \mathcal{K}}[\text{Enc}(k, m) = c] = 1/26$. Therefore it follows that for every $m_1, m_2 \in \{A, \dots, Z\}$ we have

$$\text{Prob}_{k \in \mathcal{K}}[\text{Enc}(k, m_1) = c] = 1/26 = \text{Prob}_{k \in \mathcal{K}}[\text{Enc}(k, m_2) = c]$$

However, for $\ell \geq 2$ the perfect secrecy requirement is not met. Take $\ell = 2$, $m_1 = AA$, $m_2 = AB$, and $c = CC$. Then

$$\text{Prob}_{k \in \mathcal{K}}[\text{Enc}(k, m_1) = c] = 1/26$$

because one out of every 26 permutations k between letters maps A to C .

On the other hand,

$$\text{Prob}_{k \in \mathcal{K}}[\text{Enc}(k, m_2) = c] = 0$$

because if some permutation k maps A to C then the same permutation cannot also map B to C , and so for every key k , $\text{Enc}(k, m_2) \neq c$. \square

2 OTP cipher variations [30 points]

We showed that One-Time Pad encryption satisfies perfect secrecy if $\mathcal{M} = \mathcal{K} = \{0, 1\}^\ell$, for any ℓ . Consider some variations of the OTP cipher, where the messages and/or keys are binary strings as before but with some strings missing. Consider set \mathcal{S} of three 2-bit strings, $\mathcal{S} = \{00, 01, 10\}$.

Consider the following three variations on the OTP cipher. In all these variations the key generation algorithm chooses $k \in \mathcal{K}$ uniformly, and encryption and decryption work as in OTP, i.e. $\text{Enc}(k, m) = k \oplus m$ and $\text{Dec}(k, c) = k \oplus c$.

For each of the OTP variations below, say whether the resulting cipher is perfectly secure or not, and **prove your answer**. In each case, say what the space \mathcal{C} of the ciphertexts is, and note the *sizes* of the message space \mathcal{M} and key space \mathcal{K} . Do these sizes correlate somehow with whether or not the cipher is secure? Can you explain why?

1. Let $\mathcal{M} = \mathcal{S}^\ell$ and $\mathcal{K} = \{0, 1\}^{2\ell}$, i.e. both the message and the key are (2ℓ) -long bit strings¹ However, not every (2ℓ) -bit string can be a valid message. For example, for $\ell = 3$, we could have $m = [00, 01, 00] = 000100$ but $m = [11, 00, 11] = 110011$ is not in \mathcal{M} because $11 \notin \mathcal{S}$.

Solution: Note that $|\mathcal{M}| = 3^\ell$ and $|\mathcal{K}| = 4^\ell$ so in this version of the OTP cipher we have actually more keys than messages. Therefore this cipher could in principle be perfectly secure. Note that $\mathcal{C} = \{0, 1\}^{2\ell}$. Note also that for every 2-bit message fragment $m' \in \mathcal{S}$ of any message $m \in \mathcal{M}$ and every 2-bit ciphertext fragment $c' \in \{0, 1\}^2$ of any ciphertext $c \in \{0, 1\}^{2\ell}$, there is a unique 2-bit fragment of the key $k' = c' \oplus m'$ which maps m' into c' . Therefore the same is true for the whole message m and ciphertext c : For all $(m, c) \in (\mathcal{M}, \mathcal{C})$ there is a unique $k = m \oplus c$ which maps m into c . Therefore probability $\text{Prob}_{k \in \mathcal{K}}[\text{Enc}(m, k) = c]$ is the same for every message m , and hence perfect secrecy follows. \square

2. Let $\mathcal{M} = \{0, 1\}^{2\ell}$ and $\mathcal{K} = \mathcal{S}^\ell$

Solution: Here $|\mathcal{M}| = 4^\ell$ and $|\mathcal{K}| = 3^\ell$, so the fundamental theorem from the 1st lecture, which says that a cipher can be perfectly secure only if $|\mathcal{K}| \geq |\mathcal{M}|$ implies that this cipher *cannot* be perfectly secure. That would be a good-enough answer, but we can also convince ourselves of this on some specific example. Note first that $\mathcal{C} = \{0, 1\}^{2\ell}$. Take $\ell = 1$, $m = 00$, and $c = 11$. To map m into c we'd need a one-way pad $k = m \oplus c = 11$, but $11 \notin \mathcal{S}$, so $\text{Prob}_{k \in \mathcal{K}}[\text{Enc}(k, 00) = 11] = 0$, while, for example, $\text{Prob}_{k \in \mathcal{K}}[\text{Enc}(k, 11) = 11] = 1/3$ because $k = 00$ can be used to map m into c . \square

3. Let $\mathcal{M} = \mathcal{K} = \mathcal{S}^\ell$. *[[Hint: This one is actually not perfectly secure...]]*

Solution: Here $|\mathcal{M}| = |\mathcal{K}| = 3^\ell$, so the fundamental theorem from the 1st lecture says that this cipher in principle could be perfectly secure. That's why we said that this one is tricky, because nevertheless this cipher actually fails to satisfy the perfect secrecy requirement.

The reason is that there are ciphertexts which can correspond only to some, but not all, plaintexts. And therefore the ciphertexts do carry information about the plaintext. Take $\ell = 1$, in which case $\mathcal{C} = \{0, 1\}^2$. Take for example $c = 11$, and note that it can correspond to $m_1 = 10$ if $k = 01$. However $c = 11$ cannot correspond to $m_2 = 00$ because then $k = c \oplus m$ would have to be 11 , which is not a valid key. Therefore triple $(m_1, m_2, c) = (10, 00, 11)$ violates the perfect secrecy requirement. \square

¹We use notation \mathcal{A}^n to denote a set of n -long sequences $[A_1, A_2, \dots, A_n]$ where each A_i is an element of \mathcal{A} . Using this notation, $\{0, 1\}^n$ denotes a set of all n -long binary strings.