Problem Set 1

We will work on the following exercises together during the work session on Wednesday, 2 September 2020.

The problems marked with ★ are more challenging bonus problems, often referring to extra material covered only in the [KL] book (or elsewhere), but possibly not in the videos. These bonus problems are optional and will not be part of the material you have to master for the final exam. All students should be able to solve all non-starred problems.

Problem 1: Formal definitions

In the Katz-Lindell [KL] book, the *shift cipher* is informally described as follows: "Algorithm Gen outputs a uniform key $k \in \{0, 1, \ldots, 25\}$; algorithm Enc takes a key k and a plaintext and shifts each letter of the plaintext forward k positions (wrapping around at the end of the alphabet); and algorithm Dec takes a key k and a ciphertext and shifts every letter of the ciphertext backward k positions."

In a more formal mathematical language, we can define the shift cipher as follows: Let us identify numbers and English letters in the natural way ($a = 0, b = 1, c = 2$ etc.). Then, the message space $\mathcal{M}$ is any finite sequence $m = m_1 m_2 \cdots m_\ell$ where $m_i \in \{0, 1, \ldots, 25\}$.

$$\begin{aligned} \mathsf{Gen} &: k \leftarrow \{0, 1, \ldots, 25\}\,, \\ \mathsf{Enc}_k(m_1 m_2 \ldots m_\ell) &= c_1 c_2 \ldots c_\ell \text{ where } c_i = [(m_i + k) \mod 26]\,, \\ \mathsf{Dec}_k(c_1 c_2 \ldots c_\ell) &= m_1 m_2 \ldots m_\ell \text{ where } m_i = [(c_i - k) \mod 26]\,. \end{aligned}$$

Provide formal definitions for Gen, Enc, Dec for

(**a**) the mono-alphabetic substitution cipher,

(**b**) the Vigenère cipher.

Problem 2: Chosen-plaintext attacks

Let us consider the scenario of *chosen-plaintext attacks*. In this scenario, an attacker has the ability to obtain plaintext/ciphertext pairs (under the same unknown key) for plaintexts of its choice. For example, for the case of a shift cipher, the attacker might ask for encryptions of messages `foo` and `bar` and obtains (plaintext,ciphertext) pairs `(foo,oxx)` and `(bar,kja)`.

(**a**) Show that shift, mono-alphabetic substitution and Vigenère ciphers are all trivial to break using a chosen-plaintext attack.

(**b**) How much chosen plaintext is required to recover the key for each of the ciphers?

Problem 3: Probability theory and Bayes' rule

Let E_1 and E_2 be probability events. Then, $E_1 \wedge E_2$ denotes their conjunction, i.e. $E_1 \wedge E_2$ is the event that *both* E_1 and E_2 occur. The *conditional probability of* E_1 *given* E_2, denoted $\Pr[E_1|E_2]$ is defined as

$$\Pr[E_1|E_2] := \frac{\Pr[E_1 \wedge E_2]}{\Pr[E_2]}$$

as long as $\Pr[E_2] \neq 0$.

(**a**) Let E_1 and E_2 be probability events with $\Pr[E_2] \neq 0$. Using the definitions above, prove what is known as *Bayes' rule*:

$$\Pr[E_1|E_2] = \frac{\Pr[E_1] \cdot \Pr[E_2|E_1]}{\Pr[E_2]}\,.$$

(**b**) Let the probability that a news article contains the word *president* be 20%. The probability that it contains the word *president* if it already contains the word *Trump* is 35%. The probability that it contains the word *Trump* is 10%. Under these assumptions, what is the probability that a news article contains the word *Trump* if it already contains the word *president*?

(**c**) Let the probability that a certain cryptographic protocol is *secure* and *efficient* be 10%. The probability that it is *not secure* if it is *efficient* is 80%. What is the probability that

1. the protocol is *secure* if it is *efficient*?
2. the protocol is *efficient*?

Problem 4: Shift cipher is not perfectly secret

Let us consider the shift cipher with the following message distribution: $\Pr[M = \texttt{ik}] = 0.1, \Pr[M = \texttt{op}] = 0.3, \Pr[M = \texttt{de}] = 0.6$.

(a) Calculate the probability $\Pr[M = \texttt{ik} \mid C = \texttt{ab}]$ that the message `ik` was encrypted when ciphertext `ab` is observed.

(b) Calculate the probability $\Pr[M = \texttt{op} \mid C = \texttt{ab}]$ that the message `op` was encrypted when ciphertext `ab` is observed.

(c) Conclude that the shift cipher is not perfectly secret.

Problem 5: Perfect secrecy

Archaeologists found the following encryption table — unfortunately it is not complete.

	$>$	$\vee$	$<$	$\wedge$
$\bullet$			$<$	$\wedge$
$\leftrightarrows$	$<$		$>$	
$\circlearrowleft$	$\wedge$	$>$		
$\circlearrowright$				

On another papyrus it is explained that this encryption was used during war. The movements "left", "right", "attack" and "withdrawal" were represented by $\mathcal{M} = \{<, >, \wedge, \vee\}$. The keyspace was $\mathcal{K} = \{\bullet, \leftrightarrows, \circlearrowleft, \circlearrowright\}$. Ciphertext and plaintext used the same alphabet, i.e. $\mathcal{C} = \mathcal{M}$.

The key was picked by tossing a coin twice. For every transmission a new key was used.

Complete the table to form a perfectly secret encryption scheme.

Problem 6: Message in the clear

When using the one-time pad with the key $k = 0^\ell$, we have that $\mathsf{Enc}_k(m) = k \oplus m = m$ and the message is sent in the clear! It has therefore been suggested to modify the one-time pad by only encrypting with $k \neq 0^\ell$ (i.e., to have Gen choose k uniformly from the set of *nonzero* keys of length ℓ). Is this modified scheme still perfectly secret? Explain.

Problem 7: ★ Birthday paradox

This exercise deals with some basic probability as described in Appendix A.3 of the book [KL]. Compute the probability that there exists two students in the room that have the same birthday (ignoring the birth year).

Problem 8: ★ Asymptotic growth rate

Let ε and c be arbitrary constants such that $0 < \varepsilon < 1 < c$. Order the following terms in increasing order of their asymptotic growth rates.

$$n^n \quad \exp(\sqrt{n}) \quad 1 \quad \log\log n \quad c^{c^n} \quad n^c \quad n^\varepsilon \quad n^{\log n} \quad \log n \quad c^n$$

Hint: In some cases, it might help to express two terms you want to compare in the form $e^{\cdots}$ and then compare their exponents.

Another hint: One can prove that for any positive polynomial $p(n)$, any constant $\varepsilon > 0$ and large enough n, it holds that $p(\log(n)) < n^\varepsilon$. You may simply use this fact, or you can try to prove it.

Problem 9: ★ Logical Contrapositives

In logic, contraposition is an inference that says that a conditional statement $A \Rightarrow B$ is logically equivalent to its contrapositive $\neg B \Rightarrow \neg A$. For instance, the proposition "If the weather is good, then I'm biking to work." can be restated as "If I'm not biking to work, then the weather is bad."

State the contrapositive of the following statements:

(a) If it rains, the trees get wet.

(b) If the car drives, its fuel tank is not empty.

(c) If p is a prime, then p = 2 or p is odd.

(d) If assumption X holds, protocol Y is secure.

(e) The RSA protocol is insecure, if you can factorize efficiently.

(f) I have read the book, if I have read every sentence of it.

(**g**) If for every theorem in this book a proof is given, then it is a good math book.