**Problem Set 2** Cryptography, 6CCS3CIS
Eamonn W. Postlethwaite October 6, 2025

**Problem 1.** *On message and ciphertext spaces.*

A function $f\colon X \to Y$ is a map that associates an element $y \in Y$ to each element $x \in X$.[1] We write $f(x) = y$ to denote this association and $f(X) = \{f(x)\colon x \in X\} \subseteq Y$ for the image of $f$ in $Y$. Some properties a function may have are injectivity, surjectivity and bijectivity. We say function $f$ is

- injective when each $x \in X$ maps to a distinct $y \in Y$, i.e. $f(x) = f(x') \iff x = x'$,
- surjective when for each $y \in Y$ there exists at least one $x \in X$ that maps to $y$, i.e. $f(X) = Y$,
- bijective when it is both injective and surjective.

*i.* For each function below determine whether it is injective, surjective, bijective or none.

*a.* $f\colon \{1,2\} \to \{1,2,3\}$, $f(1) = 1$, $f(2) = 3$.
*b.* $g\colon \{1,2\} \to \{1,2,3\}$, $g(1) = 1$, $g(2) = 1$.
*c.* $h\colon \{1,2,3\} \to \{1,2\}$, $h(1) = 1$, $h(2) = 2$, $h(3) = 1$.
*d.* $i\colon \{1,2,3\} \to \{1,2,3\}$, $i(1) = 1$, $i(2) = 3$, $i(3) = 2$.
*e.* $j\colon \{1,2,3\} \to \{1,2\}$, $j(1) = 1$, $j(2) = 1$, $j(3) = 1$.

*ii.* Let $X$ and $Y$ be finite sets. Prove that if $f\colon X \to Y$ is injective then $|Y| \geq |X|$.[2] Prove that if $f$ is surjective then $|X| \geq |Y|$. Prove that if $f$ is bijective then $|X| = |Y|$.

Decryption is unambiguous when, given the secret key $sk$, it is possible to recover the encrypted message from the ciphertext. This is required for correctness. We are going to consider necessary conditions for decryption to be unambiguous. That is, given a secret key $sk$, which properties must $\mathsf{Enc}(sk, \cdot)$ have as a function to allow decryption to be unambiguous? When decryption is not unambiguous we say it is ambiguous.

Let $\mathsf{SKE} = (\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ be a secret key encryption scheme. Let the message and ciphertext space for $\mathsf{SKE}$ be $\mathcal{M}$ and $\mathcal{C}$ respectively. For a fixed secret key $sk$ let $\mathsf{E}_{sk}\colon \mathcal{M} \to \mathcal{C}$ such that $\mathsf{E}_{sk}(m) = \mathsf{Enc}(sk, m) = c$.

*iii.* Give an example of a message space $\mathcal{M}$, a ciphertext space $\mathcal{C}$ and a surjective $\mathsf{E}_{sk}$ for which decryption is ambiguous. (Hint: try to imagine a situation where, even knowing the secret key, it is impossible to know which message was encrypted.)

*iv.* Give an example of a message space $\mathcal{M}$, a ciphertext space $\mathcal{C}$ and a surjective $\mathsf{E}_{sk}$ for which decryption is unambiguous.

*v.* Prove that if decryption is unambiguous then $\mathsf{E}_{sk}$ is injective. What can you conclude about the sizes of $\mathcal{M}$ and $\mathcal{C}$ if they are both finite?

*vi.* Give an example of a message space $\mathcal{M}$, a ciphertext space $\mathcal{C}$ and an injective $\mathsf{E}_{sk}$ that is not surjective for which decryption is unambiguous. Conclude with a minimal necessary assumption for unambiguous decryption.

**Problem 2.** *On brute force attacks.*

The Caeser cipher is a historical example of an $\mathsf{SKE}$, of which more next week. Let $\mathcal{K}$ denote the key space. Let $\mathcal{M}$ and $\mathcal{C}$ denote the message and ciphertext space respectively, as above. Let $\mathcal{K} = \{0, 1, \ldots, 25\}$ and $\mathcal{M} = \mathcal{C} = \{0, 1, \ldots, 25\}^*$. Encode lowercase letters of the English alphabet as a = 0, b = 1, ..., z = 25. We define the Caesar cipher in the formalism of $\mathsf{SKE} = (\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ as

- $\mathsf{KGen}$: select $sk$ uniformly from $\mathcal{K}$,
- $\mathsf{Enc}(sk, m)$: parse $m = m_1 m_2 \cdots m_\ell$ as individual characters $m_1, \ldots, m_\ell$.
  For $i \in \{1, \ldots, \ell\}$ let $c_i = m_i + sk \bmod 26$ and return $c = c_1 c_2 \cdots c_\ell$.

[1] The symbol $\in$ is read aloud as 'in', so $x \in X$ means an element $x$ in the set $X$.
[2] For a finite set $X$ the notation $|X|$ denotes the number of elements in $X$.

- $\mathsf{Dec}(sk, c)$: parse $c = c_1 c_2 \cdots c_\ell$ as individual characters $c_1, \ldots, c_\ell$.
  For $i \in \{1, \ldots \ell\}$ let $m_i = c_i - sk \bmod 26$ and return $m = m_1 m_2 \cdots m_\ell$.

*i*. You receive the ciphertext $c = \mathrm{ck}$. If you allow acronyms and abbreviations, how many human readable messages are there that can encrypt to $c$ using the Caeser cipher, and what are the respective keys?

*ii*. Is a single ciphertext always sufficient to mount a brute force attack against an SKE?

*iii*. What is the longest ciphertext you can find that can decrypt to at least two human readable messages using the Caesar cipher?

*iv*. If an adversary has access to $\mathsf{Enc}(sk, \cdot)$ describe a (non brute force) attack against the Caesar cipher that does not need any ciphertexts.

**Problem 3.** *On Kerckhoffs' principles.*

Encryption based on codebooks can be shoehorned[3] into the formalism of an SKE.

If Alice and Bob wish to communicate they take the Oxford English Dictionary (OED)[4] and for each word in order, e.g. starting with aardvark, they choose another word, e.g. cryptography, that has not been chosen before. They each write aardvark in the left hand column of their codebook and cryptography in the same row of the right hand column. (Formally, this is a permutation of the words of the English language.) This represents key generation with the codebook being $sk$.

To encrypt a word, find it in the left hand column and write the word in the same row of the right hand column. This is easy, as hopefully the left hand column is in alphabetical order.

To decrypt a word, find it in the right hand column and write the word in the same row of the left hand column…

*i*. Consider the above with respect to each of Kerckhoffs' principles. Which are satisfied and which are not?

*ii*. How might we make decryption more efficient? (Hint: $sk$ is already huge, why not double its size?)

*iii*. How might we rewrite Kerckhoffs' principles in the modern day? Do we need to rewrite them all?

*iv*. Find an example where the failure to apply one or more of Kerckhoffs' principles has been catastrophic.

[3] This is idiomatic, it means 'forced into'.
[4] The second edition of the OED has twenty volumes and 21,730 pages.