

CS4236 Cryptography Theory and Practice Assignment 1

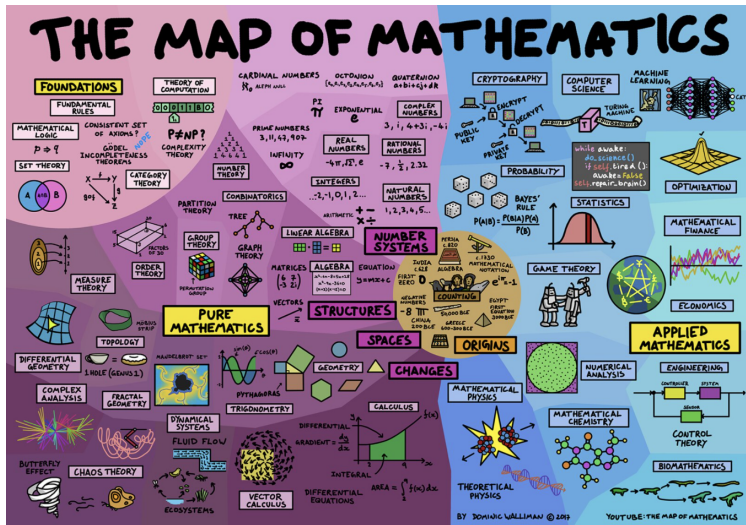
Hugh Anderson

National University of Singapore
School of Computing

August, 2022



Where are we then?



Outline

1 Admin

- Special help sessions

2 Assignment 1

- Comments on question 1
- Comments on question 2
- Comments on question 3
- Comments on question 4
- Comments on question 5

Special help sessions on Saturdays

Or extra tutorial, or open house, or town square, or ...

On Saturdays, from 2:00 to 3:00, I will run a zoom session from my home. You can join at any time, and just yell out or something (I will leave the machine running in the living room, and try to keep an eye on it).

If you have any questions, come and talk to me via zoom on Saturday:

URL: <https://nus-sg.zoom.us/j/82466546798?pwd=Z1FXZnF6OWdCQnJNeFAyTDEzKzFkZz09>

Meeting ID: 824 6654 6798

Passcode: 182428

(From session 1) Symmetric encryption:

A symmetric encryption system/scheme $(\text{Gen}, \text{Enc}, \text{Dec})$:

... comprises 3 algorithms:

- 1 $\text{Gen}(1^n)$: Typically the input is the size of the key
- 2 $\text{Enc}_k(m)$: Input is the key k and the message. For presentation, the key k is put as a subscript.
- 3 $\text{Dec}_k(c)$: Input is the key k and the ciphertext.

Note that we use this notation:

\mathcal{K} : set of all keys.

\mathcal{M} : set of all messages

\mathcal{C} : set of all ciphertexts

For correctness, we see that for all k, m , $\text{Dec}_k(\text{Enc}_k(m)) = m$.

Question 1

1. Provide a formal definition¹ of the Gen, Enc, and Dec algorithms for a ...

Comment...

The word *formal* here indicates that I am looking for a mathematical unambiguous definition. The cipher specified is a particular instance of the (abstract) symmetric encryption scheme from lecture 1 (just shown), and was presented as a triple of functions: (Gen, Enc, Dec). In that definition, the domain (keys, messages, ciphertexts) is mentioned, and also correctness.

When we talk about particular instances of schemes, the terminology we use is *Construction*, so you are being asked to write a formal definition of a *construction* for this symmetric encryption scheme. It should have

- a name (perhaps “Definition of Vigenere Encryption construction”)
- an explanation of the domains/ranges/environment (what are keys messages and ciphertexts in this case), and
- definitions of each of the three functions (Gen, Enc, Dec).

You could also give a hint as to what correctness would mean for this cipher.

¹Note that “message and plaintext” should be “ciphertext and plaintext”.

Question 1

We have seen a construction before...

In the second session there was a formal description of the one-time-pad construction (next slide). This definition had

- a name (the one-time-pad),
- an explanation of the (uniform) keys $k_i \in \mathbb{Z}_n$ and messages made up from a sequence of elements $x_i \in \mathbb{Z}_n$, but omitted to define the ciphertexts because I forgot to do that, but in any case the ciphertext is a sequence of elements $y_i \in \mathbb{Z}_n$, and it had
- definitions of the three functions (Gen, Enc, Dec).

There was a hint as to what correctness would mean for this cipher.

Q1, assignment 1...

I expect something like the formal definition, but for Vigenere. It might be better to give the name, explain in a short paragraph your keys, messages and ciphertexts, and then define the three functions.

By the way, in my course slides, to keep them from being too dense, I just put the main elements of definitions. The textbook is a bit more precise - see definition 2.8 on the top of page 33. I expect you to be precise.

(From session 2) Construction: One time pad

Formal definition: $(\text{Gen}, \text{Enc}, \text{Dec})$

$\text{Gen}(1^n)$: The key is a random sequence of elements in \mathbb{Z}_n :
 $K = k_1, k_2, k_3, \dots$

$\text{Enc}_k(X)$: The plaintext is a sequence of elements in \mathbb{Z}_n ,
 $X = x_1, x_2, x_3, \dots$

The ciphertext is

$$Y = \text{Enc}_k(X) = (k_1 + x_1) \bmod n, (k_2 + x_2) \bmod n, \dots$$

$\text{Dec}_k(Y)$: To decrypt the ciphertext $Y = y_1, y_2, y_3, \dots$, the plaintext is
 $X = \text{Dec}_k(Y) = (y_1 - k_1) \bmod n, (y_2 - k_2) \bmod n, \dots$

Commentary

Correctness: Note that $x_i = ((x_i + k_i) \bmod n) - k_i \bmod n$

When $n = 2$, then the key is a sequence of binary bits, and encryption (and decryption) is equivalent to *xor'ing* the key. If the length of the (random) bitstring is ℓ then the probability of $X = Y$ is $2^{-\ell}$.

Question 2

2. Prove the correctness of the cipher...

Comment...

In question 2 you are asked to prove the correctness of your construction. As we were just reminded, correctness of a symmetric scheme is shown by showing that (for all keys k and messages m) if you encrypt the message, and then decrypt the result, you should end up with your original message. i.e.

$$\forall k, m : \text{Dec}_k(\text{Enc}_k(m)) = m$$

Your task in question 2 is to prove this using *your* definitions of $\text{Enc}_k(m)$ and $\text{Dec}_k(c)$.

I expect this to be done using the style of proof outlined in class in the second lecture. If you rely on some mathematical identity you should state what it is in the steps.

Question 3

3. Explain why the cipher is poorly constructed...

Comment...

In question 3 you are asked to think about what might be the weakness of this cipher.

In general, a cipher like this could be perfectly secret, as long as the keys are random, and the repeated key length was as long as, or longer than, the message^a. However, in this case the construction is slightly different, and even if $t \geq |M|$, this difference leads to a weakness.

In your explanation, make sure I understand why you think it is poorly constructed, perhaps by showing an attack.

^aNote that in class, some time was spent discussing the one-time-pad, and pointing out that XOR is just modulo-2, and that these schemes work modulo-anything. Note also that the one-time-pad definition (Construction 2.8, pg33) in the textbook is for bitwise (XOR) one-time pads, whereas the definition I gave for the one-time-pad is for modulo-anything.

Question 4

4. Find the following, clearly stating *why* for each...

Comment...

This question just asks you to do some worked probability examples, to confirm you can use the probability rules and theorems given in class.

Some of the questions might be a bit tricky, some of them might rely on the key and plaintexts being independent (they normally are), and some may require you to use the rules and theorem shown in class. Finally I think that \wedge associates more strongly than $|$, so $E_1, E_2|E_3$ is $(E_1, E_2)|E_3$.

Each answer should be accompanied by a brief explanation, or reference to the particular rule(s) exploited in giving the answer.

Question 5

5. Prove or refute: ...

Comment...

Here we have a question which asks you to prove, or disprove, another possible variation of a definition of *perfect secrecy*.

We saw in session 2 some variations on the definition of perfect secrecy. In particular the slides about definitions 2.3 and 2.3(a) are alternative definitions. Outlines of proofs to show the equivalence were given.







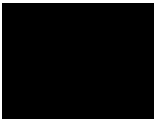
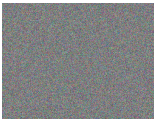

To answer Q5, you could try to see if you can prove that this definition implies 2.3, or 2.3(a) and/or if 2.3 or 2.3(a) implies this definition. In each case you can always write down the first and last line of the proof. If you succeed in both proofs then you have proved this.

Alternatively you may find one or other of the proofs does not seem to work. If you can find a reason for the failure, you have disproved, or refuted the equivalence of the LHS and RHS of the equation.

In either case, you start by exploring the equation.

More context for question 5

A reminder on one aspect of perfect secrecy

m	\oplus	k	\rightarrow	c
	\oplus		\rightarrow	
	\oplus		\rightarrow	
	\oplus		\rightarrow	

Perfect secrecy is not dependent on uniformly distributed messages.