CS6111: Foundations of Cryptography

Roll No: CS16B039

Instructions

- Deadline is Monday, August 19.
- We encourage submissions by Latex. Paper is also accepted.

References

- Introduction to Cryptography - Delfs and Knebl
- A Graduate Course in Applied Cryptography - Boneh and Shoup (link)
- Introduction to Modern Cryptography - Katz and Lindell
- Cryptography: Theory and Practice - Douglas Stinson (3rd edition)
- Handout 1 - Entropy Axioms
  - It gives an overview of entropy with many definitions. Page 7 starts talking about the axioms.

1 Perfect Secrecy

1. (8 points) We defined Shannon's entropy in class and mentioned its uniqueness. "Handout1 - Entropy Axioms" shows how the uniqueness of the definition can be derived from axioms.
   (a) (2 points) Explain, with examples of distributions, how Shannon's entropy captures the amount of information in bits we obtain after an event.
   (b) (2 points) Define conditional entropy and relative entropy. Justify why their definitions make sense.
   (c) (4 points) Justify both sets of **axioms** for Shannon's entropy (pg 8,9). Explain how you believe they each capture the essential properties of entropy we require.

Solution:
(a) Every 'finite schema' (or) a 'complete system of events' has an associated probability distribution that corresponds to the probabilities of different ways in which the event can manifest. Before the event could occur, if an observer is aware of the probability distribution of the manifestations, he can bet on the probable outcome based on it.
In this kind of bet, there is a risk involved which corresponds to the inherent uncertainty of the distribution.

Shannon's Entropy is a way of quantifying this uncertainty or risk and is given by:

$$H(\mathbf{p}) = -\Sigma_{i=0}^{k} p(e_k) log p(e_k)$$

where e_k is the probability of k^{th} manifestation occuring.

Let us understand this better with examples.

Example 1: Consider a group of 8 men with one of them being the culprit of a criminal case. There is a witness that can identify the culprit correctly. However, the witness can only answer 'Yes' or 'No' to any question he is asked. We can consider that with each answer, **the witness reveals one bit of information** given this binary scheme. Thus, to crack this case, one needs to ask the question

"Is he in the right half or the left half?"

at least thrice to spot the culprit. This means that once this finite scheme of probable culprits has been solved, it revealed information equivalent to asking three 'Yes/No' questions. Thus the event has inherent uncertainty (or the information revealed when the event actually occurs) of 3 bits.

Let us look at the Shannon's entropy. As no previous information is known about any of the accused, the distribution would be uniform. i.e

$$P = (\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, ..., \frac{1}{8})$$

Calculating the Shannon's Entropy for this:

$$8 * 1/8 * log(8) = 3$$

Which also turns out to be 3. **Example 2:** Let us consider an **non-uniform distribution** and see how we can look at uncertainty in terms of number of 'Yes/No' questions. Suppose there is a course of 9 people, one of them is auditing the course. The professor wants to find out who it is by asking the TA who can answer only a 'Yes' or a 'No'. The professor always finds one of the students in several classes and has the following probability distribution in his mind:

$$\mathbf{p} = (1/2, 1/16, 1/16, 1/16, 1/16, 1/16, 1/16, 1/16, 1/16)$$

which is to say: He knows that one of the students is very likely to be the auditing student but is equally unsure about the rest of the students. We can measure this uncertainty by judging the **expected optimal number of 'Yes/No' questions** he asks his TA to find the auditing student.

Since he knows that one of the students is most likely, he asks about him the first. Half of the times, he actually is the student auditing. In half of the cases, when he is not the one auditing, he takes at least three **additional** questions to find out the student auditing for the course.

$$E('Y/N' Questions) = (1/2) * (1) + (1/2) * (1 + 3) = 2.5$$

Now let us look at the Shannon's entropy for this case which is:

$$1/2 * log(2) + 8 * 1/16 * log(16) = 2.5$$

From these two examples, we see that, **the uncertainty as quantified by Shannon's entropy indeed matches with the commonsense notion of uncertainty** as depicted by 'Yes/No' paradigm.

(b) **Conditional entropy** is defined by:

$$H(X/Y) = -\Sigma\Sigma p(y)p(x/y)logp(x/y)$$

Justification: Suppose there exists two finite scheme of events **X** and **Y**, then the conditional entropy can be conceptually seen as the **expected information gained about X when the event Y happens** and the result is known. We can prove this by considering a row $(Y = y_i)$ in the X*Y matrix and considering the relative probabilities in each column:

$$\text{Expected gain in information about X given Y} = -\Sigma p(y) * (\Sigma p(x/y)logp(x/y))$$

Keeping in mind that X*Y form a discrete matrix of finite compound events, we can simply refactor the above euqation as :
Expected gain in information about X given Y $= -\Sigma\Sigma p(y)p(x/y)logp(x/y)$
This matches the definition of conditional entropy.

Relative Entropy:
Relative Entropy, often also called as **K-L Divergence** is a measure of how close two given probability distributions are.
It is given by:

$$H(p|q) = \Sigma_{j=1}^{n} p_j log(\frac{p_j}{q_j})$$

Justification: We have seen that conditional entropy is the information leaked about X when Y's event is captured. Unlike conditional entropy, relative entropy does not capture a notion of common information. Instead it only talks about the closeness of two distribution in a mathematical sense.

(c)The Shannon-Khinchin Axiom is a compound set of three axioms. Let us call them (1), (2), (3). The Faddeev axiom is a compound axiom of [1], [2], [3] axioms. Although these axioms look different, they are essentially equivalent. To understand, this let us compare (1) and [1]:

[1] is very simple and does not capture much information except that Shannon's entropy is positive at at least on point in [0,1] and that it is a continuous function. On the other hand, (1) has more information saying these things: At each point Shannon's entropy is positive. It is continuous. It has it maxima at uniform distributions.

[2] is just saying that the permutation of possible manifestations of an even in the probability distribution should not affect to resulting Shannon's entropy mathematically. On the other hand, (2) talks about how impossible manifestations of an event (i.e probability of zero) should not affect the Shannon's entropy.

[3] looks simpler than (3). However, it is just as powerful. This is because there is a finite-step reduction

possible from [3] to arrive at (3). Both these axioms are talk about the inductive generation of Shannon's entropy. [3] can be looked at as: decomposing a compound event into two subevents, whereas (3) does this at a scale of multiple decompositions of multiple events.

(1),(2) and [1],[2] can be seen as boundary or baseline conditions that regularize the Shannon's entropy to fall into the co-domain in a structured way. Adding to this, [3],(3) serve to solidify this structured mapping into a valid function defined from $[0,1]^n$ to $\mathbb{R}$. But the semantic information captured by the two seemingly different axioms is essentially the same.

2. (4 points) Suppose a cryptosystem (E, D) achieves perfect secrecy for a particular distribution $P = (p_1, \ldots, p_n)$ on a plaintext set $M = \{m_1, \ldots, m_n\}$. Prove that the system is also perfectly secret for all other probability distributions P' on M.

Solution: The key to solve this problem is to consider the conditional probability $p(c_j/m_i)$. We find that this is same as $p(E(m_i) = c_j)$.
Claim 1: $\forall i, j\ Pr(c_j/m_i)$ is an invariant across all distributions of P on M.
Proof: We have $p(c_j/m_i) = p(E(m_i) = c_j)$. Since E is the encryption algorithm of the given crpyptosystem, we can consider it as an probabilistic randomized algorithm and it's output distribution solely depends on the plaintext m_i and the coin tosses that occur during the execution of the algorithm. Therefore, it is immaterial to E, the distribution from which the given m_i is picked from and hence $\forall i, jPr(c_j/m_i)$ is an invariant across all distributions of P on M.

From proposition 9.4 [Delfs 3rd ed.] we have:
Proposition: For an encryption E, the following statements are equivalent:
1) E is perfectly secret
2) $Pr(c/m_i) = Pr(c/m_j)\ \forall i, j < |M|$
Let us make use of this proposition to prove that the given cryptosystem (E,D) is perfectly secret for all distributions over M. It is given that perfect secrecy is achieved for a particular distribution

$$\mathrm{P} = (p_1, p_2, p_3, \ldots\ldots, p_n)$$

From the above proposition, over this distribution the following holds true:

$$Pr(c/m_i) = Pr(c/m_j)\ \forall i, j < |M| \text{ over distribution P}$$

From claim 1, we have that $\forall i, j\ Pr(c_j/m_i)$ is an invariant across all distributions. This means for an arbitrary new distribution $P^{'}$ we still have:

$$Pr(c/m_i) = Pr(c/m_j)\ \forall i, j < |M| \text{ over distribution } P^{'}$$

So for a given distribution $P^{'}$, from the above proposition we have that E is perfectly secret.
As our choice of $P^{'}$ has been arbitrary, we can generalize this result over all possible distributions over M.
Therefore, the given cryptosystem achieves perfect secrecy over all plaintext distributions.

3. (4 points) What is the affine cipher? Show that the number of keys (a, b) when $m = 26$ is 312.

(a) (2 points) Prove prefect secrecy when each key is chosen with equal probability.

(b) (2 points) Suppose instead that we choose (only valid values of) a according to a given probability distribution P and b is chosen randomly. Prove that perfect secrecy still holds.

Solution: As m i.e $|M|$ is 26, without loss of generality, let us assume that the message space is equivalent to the English Alphabet and let us denote it as Σ.

Affine cipher is a monoalphabetic substitution cipher that substitutes an alphabet (just a message in our case) for another alphabet (another message). To do this, it maps each alphabet to a integer between 0 and 25. In our case, A maps to 0, B maps to 1, ..., Z maps to 25.

Affine cipher uses an ordered pair of integers as its key. Each of these integers also comes from the set {0,1,2,.....,25}.

Working of Affine Cipher: The affine cipher considers each alphabet and muliplies it with the left key in the modulo 26 space. Next it shifts i.e adds the right key to the resulting value and considers modulo 26 again. i.e

$$\text{Affine}(X,(a,b)) \mapsto (aX+b) mod N$$

During the decryption, the ciphertext is taken and we add 'N-b' to it. Then we multiply it with modular inverse of a to extract the actual message back. i.e

$$\text{Decrypt}(X,(a,b)) \mapsto a^{-1}(X+N-b) mod N$$

However, as we know from the theory of modular inverses, a^{-1} might not always exist. Infact, :

$$\forall a \exists x, xa \equiv ax \equiv 1 \text{ (mod N) if and only if gcd(a,N)=1}$$

Therefore, there is a constraint on selection of 'a' where as 'b' can come from any value in {0,1,2,.....,25}. The possible values of a that satisfy gcd(a,26)=1 between 0 and 25 are:

1,3,5,7,9,11,15,17,19,21,23,25

Therefore the number of valid 'a's are: 12

The number of valid 'b's are: 26

Therefore, the number of valid keypairs (a,b) are: 26*12 = 312

Perfect secrecy:

(a) We know that for each 'a' there exist 26 possibilities for 'b'. It is given that the key-pairs are chosen at random. Let us look at how this logically would mean perfect secrecy.

Observation: If two distinct keypairs have the same value for 'a', they cannot map a message to the same ciphertext. i.e if $b_1 \neq b_2$, then Aff(m,(a,b_1)) $\neq$ Aff(m,(a,b_2))

Corollary1: If the choice of 'a' is fixed, there exists exactly one choice of 'b' mapping a given alphabet to another given alphabet. As the keys are chosen at random,

$$Pr(b_i/a_j) = 1/26 \text{ for all i,j } \epsilon \{0,1,2,....,25\}x\{0,1,2,....,11\}$$

We can write the following equation from additive law of conditional probabilites:

$$\forall j,\ \forall m \epsilon M,\ Pr(\tfrac{c_j}{m}) = \Sigma_{i=1}^{12} Pr(E(a_i,b,m)=c_j) * Pr(a=a_i)$$

From corollary1, that this is equvivalent to:

$$\forall j,\ \forall m \epsilon M,\ Pr(\tfrac{c_j}{m}) = \Sigma_{i=1}^{12} Pr(\tfrac{b=b^*}{a=a_i}) * Pr(a=a_i)$$

where b^* is the unique b from corollary1.

For a fixed a, as the choice of b's come randomly, we have:

$$\forall j,\ \forall m \epsilon M,\ Pr(\tfrac{c_j}{m}) = \Sigma_{i=1}^{12} \tfrac{1}{26} * Pr(a = a_i)$$

Similarly, considering the keys are chosen at uniformly random then the value of 'a' also comes uniformly from the set: 1,3,5,7.....,25. :

$$\forall j,\ \forall m \epsilon M,\ Pr(\tfrac{c_j}{m}) = \Sigma_{i=1}^{12} \tfrac{1}{26} * \tfrac{1}{12} = \tfrac{1}{26}$$

Using proposition 9.4, [Delfs 3rd ed.], this definition is equivalent to the original definition of **perfect secrecy** (as we have shown that Pr(c/m1) = Pr(c/m2) for all m1,m2,c)

(b) The argument is similar even for the case of non-uniform distribution of 'a'. With the same chain of reasoning, we can reach the step:

$$\forall j,\ \forall m \epsilon M,\ Pr(\tfrac{c_j}{m}) = \Sigma_{i=1}^{12} \tfrac{1}{26} * Pr(a = a_i)$$

Now, we take 1/26 out of the fraction and the result is:

$$\forall j,\ \forall m \epsilon M,\ Pr(\tfrac{c_j}{m}) = \tfrac{1}{26} * \Sigma_{i=1}^{12} Pr(a = a_i)$$

$$\forall j,\ \forall m \epsilon M,\ Pr(\tfrac{c_j}{m}) = \tfrac{1}{26}$$

(because $\Sigma_{i=1}^{12} Pr(a = a_i) = 1$)

Using proposition 9.4, [Delfs 3rd ed.], this definition is equivalent to the original definition of **perfect secrecy** (as we have shown that Pr(c/m1) = Pr(c/m2) for all m1,m2,c)

4. (4 points) Let (Enc, Dec) be a perfectly-secure scheme. Define a new scheme $Enc'((k_1, k_2), m) = (Enc(k_1, k_2), Enc(k_2, m))$. Prove that Enc' is perfectly-secure.

Solution: We see that $Enc^{'}$ encrypts the plaintext message producing an ordered pair as the ciphertext. Let us arbitrarily choose an ordered pair in the CxC space and call it (a,b).

Let us look at the following probabilities:

$$Pr(Enc^{'}(m_1) = (a, b)) \text{ and } Pr(Enc^{'}(m_2) = (a, b))$$

To prove that $Enc^{'}$ is perfectly secure, we need to prove that these two probabilities are the same for an arbitrary pair of m_1 and m_2 from proposition 9.4, [Delfs 3rd ed.]. (Note a and b are already considered in an arbitrary fashion).

Claim 1: $Pr(Enc^{'}(m_1) = (a, b)) = Pr(Enc^{'}(m_2) = (a, b))$

To prove claim 1, we will first decompose each of the above probabilities as:

$$Pr(Enc^{'}(m_i) = (a, b)) \ = Pr(Enc(k_1, k_2) = a, Enc(k_2, m_i) = b)$$

This is from the definition of $Enc^{'}$, which is:

$$Enc'((k_1, k_2), m) = (Enc(k_1, k_2), Enc(k_2, m))$$

We will revisit claim 1 and prove it after proving claim 2.

Claim 2: For any given m_i the events: $Enc(k_1, k_2) = a$, $Enc(k_2, m_i) = b$ are independent.

Proof: Consider the events $Enc(k_1, k_2) = a$, $Enc(k_2, m_i) = b$. These two events correspond to two distinct runs of the encryption algorithm 'Enc'. Enc's output depends only on: key, plaintext message, coin tosses during the execution. The coin tosses in the two runs happen independently. Also the keys are chosen uniformly and independent of any other random variable. Therefore, the only possibility of dependence would arise from the using key of the latter event as the plaintext message in the former event.

Let us argue that using key of one run of 'Enc' as plaintext message of another still holds these algorithm runs independent. To prove this, consider that 'k' must have been randomly chosen from the keyspace. As encryption algorithm runs on both K as well as M space, let us assume that these spaces are same. i.e K = M = $\{0,1\}^n$. Now since we know that $\forall m \epsilon M,\ Pr(C = a/M = m)$ is same for Enc Algorithm. This implies, if $Pr(Enc(k_1, k_2) = a, Enc(k_2, m_i) = b) > Pr(Enc(k_1, k_2) = a) * Pr(Enc(k_2, m_i) = b)$, then it must symmetrically hold for all the key choices from K based on the invariant from last sentence. i.e

$$\forall k,\ Pr(Enc(k_1, k) = a, Enc(k, m_i) = b) > Pr(Enc(k_1, k) = a) * Pr(Enc(k, m_i) = b)$$

Consider the summation over K for this, will give an absurd result that the occurrence of (a,b)'s for a fixed b can happen more often than b's occurance on the LHS of the ordered pair.

Therefore $Enc(k_1, k_2) = a$, $Enc(k_2, m_i) = b$ are independent.

Now let us revisit claim 1. From claim 2, we have:

$$Pr(Enc^{'}(m_i) = (a, b)) \ = Pr(Enc(k_1, k_2) = a) * Pr(Enc(k_2, m_i) = b)$$

As Enc is perfectly secret, we have the invariant:

$$Pr(Enc(k_2, m_i) = b) = Pr(b)$$

Therefore from the above invariant: $Pr(Enc^{'}(m_1) = (a, b)) = Pr(Enc^{'}(m_2) = (a, b))$
Using proposition 9.4, [Delfs 3rd ed.], this definition is equivalent to the original definition of **perfect secrecy** (as we have shown that Pr(c/m1) = Pr(c/m2) for all m1,m2,c).

Variations to Perfect Security

Let (Enc, Dec) be an encryption scheme on (M, K, C).

Indistinguishability Attack Game - Experiment b:

1. Attacker Eve chooses $m_0, m_1 \in M$.
2. Challenger chooses a random $k \in K$ and gives Eve $c = Enc_k(m_b)$.
3. Eve outputs her guess b'.

Eve wins if $b' = b$.

Advantage: Let W_b be the probability that Eve outputs 1 in Experiment b. The semantic-security advantage of the attacker is defined as $\mathsf{SSAdv}(Eve, Enc) = |W_0 - W_1|$.

- If the scheme is perfectly secure, then the probability that Eve wins is $1/2$. Equivalently, Eve's advantage is 0.

5. (4 points) Consider a variant of the one-time pad where $M = C = \{0,1\}^L$ but the keyspace $K \subset \{0,1\}^L$ is restricted to strings with an even number of 1s. Prove that this version is not perfectly secure and an adversary can win the indistinguishability game with probability 1.

Solution: The given attack is a 'Chosen Plaintext attack'. Consider two possibilities:

Possibility1 : L is odd

Then Eve can choose: m_0 as 0000.....0 and m_1 as 1111.....1. As the keys are restricted to contain only even number of 1's, XORing with m_0 always gives even number of 1's whereas XORing with m_1 always gives odd number of 1's. Clearly they are distinct after encryption using any key in keyspace. Thus Eve can win the indistinguishability game with probability 1 using the following strategy:

Choose $\delta = 0$ if c, the ciphertext has even number of 1's,

Choose $\delta = 1$ otherwise

Possibility2 : L is even

In this case, Eve can choose: m_0 as 0000.....0 and m_1 as 1111.....10. As the keys are restricted to contain only even number of 1's, XORing with m_0 always gives even number of 1's whereas XORing with m_1 always gives odd number of 1's. Clearly they are distinct after encryption using any key in keyspace. Thus Eve can win the indistinguishability game with probability 1 using the following strategy:

Choose $\delta = 0$ if c, the ciphertext has even number of 1's,

Choose $\delta = 1$ otherwise

Clearly, with this strategy, Eve can win with a probability of 1.

6. (4 points) Define an encryption scheme to be ϵ-perfectly secure if the probability that Eve wins the game is $1/2 + \epsilon$. Show that ϵ-perfect secrecy can be achieved with $|K| < |M|$ when $\epsilon > 0$. Prove a lower bound on the size of K in terms of M, ϵ.

Solution:

First, let us construct an ϵ-perfect scheme with $|K| < |M|$. Later we will argue that $|K| < (1-\epsilon)|M|$ for any epsilon perfect scheme.

Example:

Consider K = {0,1} and M = {00,01,10,11}. Then we can find from M x C table that Eve can win the game with an advantage of $\frac{1}{2}$. This is because she can choose 00, 01 which will always produce different ciphertexts. Therefore, such case of relaxed perfect secrecy is possible with, $|K| < |M|$

Proof of lower bound: Suppose that$|K| < (1-\epsilon)|M|$. Let us consider an uniform distribution over the plaintext message space to arrive at an contradiction.

Let 'c' be a member of ciphertext space that occurs at least once. Decrypting 'c' with multiple keys would give multiple plaintexts. Let us call the set of all plaintexts retrieved when 'c' is decrypted with all possible keys as Dec(c) i.e

$$Dec(c) = \{m|\ \exists k Dercypt(c,k) = m\}$$

Clearly,$|Dec(c)| \leq |K|$. As we have assumed that $|K| < (1-\epsilon)|M|$, we can see that $|Dec(c)| < (1-\epsilon)|M|$. This means that at least ϵ fraction of M is not in Dec(c). (Notice the strict inequality above). i.e these have the probability $Pr(\frac{M=m^{'}}{C=c}) = 0$.

Mathematically, we can write them to get the inequality as:

$$Pr(m \notin Dec(c)) \geq \epsilon$$

And for a single such m, say $m^{'}$,

$$Pr(\frac{M = m^{'}}{C = c}) = 0$$

Combining them, we have:

$$|Pr(\frac{M = m^{'}}{C = c}) - Pr(M = m^{'})| \geq \epsilon$$

This contradicts the ϵ perfect secrecy, because now we can design a strategy to get an advantage of over ϵ for Eve.

Contradicting Strategy:

In the indistinguishability game, let Eve choose two messages m^1 and m^2 randomly. By previous analysis, if Eve is aware of 'Void plaintext subsets' (the plaintexts a given c can never map to) for all 'c's, then he can win with a proabability of 1 whenever exactly one message comes in this void region. Note that both messages cannot come from a void region as m corresponds to encryption of one of these two messages. Taking these into account, his winning probability is atleast:

$$\frac{2\epsilon(1-\epsilon)(1) + (1-2\epsilon(1-\epsilon))(1/2)}{1-\epsilon^2}$$

$$= \frac{\frac{1}{2} + \epsilon - \epsilon^2}{1-\epsilon^2}$$

Given that $\epsilon < 1/2$. We can see that this fraction is$> \frac{1}{2} + \epsilon$ which contradicts the **ϵ-perfect secrecy assumption.**
Therefore, $|K| \geq (1 - \epsilon)|M|$

7. (4 points) Assume that we only require an encryption scheme Enc,Dec to satisfy the following: For all $m \in M$ we have $\mathsf{Pr}[Dec_k(Enc_k(m)) = m] \geq 2^{-t}$ (for some parameter t). The probability is taken over the choice of key and randomness used in the encryption algorithm. Show that perfect secrecy can now be achieved with $|K| < |M|$ when $t \geq 1$. Prove a lower bound on the size of K in terms of M and t.

Solution: We will first show that the such Enc, Dec scheme can be achieved with $|K| < |M|$ by constructing an example. Later, we will argue a proof for lower bound of K in terms of M and t.

Example
Consider M = $\{0,1\}^n$ and K = $\{0,1\}^{n-t}$, with Enc(m)= $[m]_{1,n-t} \oplus k$, i.e. the xor of the first (nt) bits with key bits. Here we can see that length of c is less than length of m.
This encryption scheme is perfectly secret. The best Eve can do here, in terms of correctness of m, for the last 't' bits of m is to guess them. This is because there is no possibility of achieving a probability $> 2^{-t}$. The rest of it is very similar to Vernam's One-Time Pad. i.e

$$Pr(\frac{C = c}{M = m_1}) = Pr(\frac{C = c}{M = m_2})$$

For an arbitrary m1, m2 pair. This is because, for a any given prefixes, say p1,p2 here, the Vernam's pad perfect secrecy assumption holds. This means that for any two messages, thus any two arbitrary prefixes, the above probabilites are the same.
At the decryption side, we see that the ciphertext can be inverted to only get first n-t bits of the plaintext message. However, the relaxed condition on decryption probability makes it possible to do away by just guessing the last 2 bits. Therefore, assuming t > 0, **we have an perfectly secret encryption scheme with** $|K| < |M|$.

Proof of lower bound
Claim $|K| \geq |M| * 2^{-t}$
To arrive at a tight inequality, we need to assume that for at least one of the ciphertexts, say 'c', the equality $\mathsf{Pr}[Dec_k(Enc_k(m)) = m] = 2^{-t}$ holds. Otherwise we can always define a new 't' such that this happens.
Consider that ciphertext 'c'. ecrypting 'c' with multiple keys would give multiple plaintexts. Let us call the set of all plaintexts retrieved when 'c' is decrypted (including guessing) with all possible keys as Dec(c) i.e

$$Dec(c) = \{m | \exists k Dercypt(c, k) = m\}$$

Now, suppose the contrary of our claim. i.e $|K| < |M| * 2^{-t}$ Then, from regular encryption scheme

assumption of $|C| = |K|$, we have the following:

$$|C| < |M| * 2^{-t}$$

(Note the strict inequality) This means that, if the encryption scheme is visualized as a relation, one of the c's must have at least 2^t incoming edges. i.e one of the c's have more than 2^t messages associated with it.

From the perfect secrecy, as we have Pr(m/c)=Pr(m). Therefore all of these messages are equally likely for 'c'. This removes the possibility that there is a decryption of this 'c' that can achieve an accuracy of over 2^{-t}. A contradiction to existence of such encryption scheme with perfect secrecy. **Therefore,** $|K| \geq |M| * 2^{-t}$

Let (Enc, Dec) be an encryption scheme on (M, K, C).

Parity Prediction Game:

1. Challenger chooses a random $m \in M$ and $k \in K$ and gives Eve $c = Enc_k(m)$.
2. Eve guesses $parity(m)$ and wins if she is correct.

Advantage: Let W be the probability that Eve outputs the correct parity. The advantage of the attacker is defined as $\mathsf{ParityAdv}(Eve, Enc) = |W - 1/2|$.

8. (4 points) Consider the above Parity Prediction Game. Prove, by reduction, that a semantically secure scheme is also secure against parity prediction: that is, show that if there exists an attacker with advantage ϵ in predicting parity, then we can get advantage 2ϵ in the Indistinguishability game.

Solution: Let us name the Parity Predictor (with advantage ϵ) 'Paris'. We will form a strategy for Eve in such a way that she can obtain 2ϵ advantage in the Indistinguishability game.
When Eve plays the indistinguishability game, as she can choose the plaintext messages, let her choose 1000.....00 , 00000.....00 as m_0 and m_1 respectively. Notice that the chose plaintexts have different parities. When she receives a ciphertext, say c, let Eve communicate this 'c' to Paris 7 times. Suppose x times Paris says that the corresponding plaintext message has a parity of 0 and y times that parity is 1, we have:

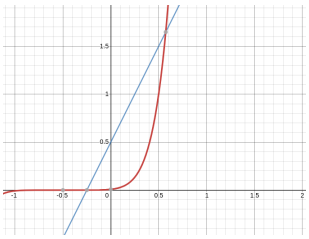
$$\text{x+y} = 7 \text{ and } \text{x} \neq \text{y} \text{ as 7 is odd}$$

As x $\neq$ y, one of them is greater than the other. Eve can then choose the following strategy:

$$\text{Choose } \delta = 0 \text{ if x} > \text{y and choose } \delta = 1 \text{ otherwise}$$

The probability with which Eve will be correct matches with the probability that Paris will be correct in at least in four of seven trials. Which is:
Pr(Eve is correct) = Pr(Paris is correct ≥ 4 times) $> (1/2+\epsilon)^7 + \binom{7}{1}(1/2+\epsilon)^6(1/2-\epsilon) + \binom{7}{2}(1/2+\epsilon)^5(1/2-\epsilon)^2 + \binom{7}{3}(1/2+\epsilon)^4(1/2-\epsilon)^3$

We know that advantage i.e $\epsilon < 1/2$ always. Therefore the last three terms are always positive. Let us now consider only $(1/2 + \epsilon)^7$.



From the graph in the above figure, clearly $(1/2+\epsilon)^7 > (1/2+2\epsilon)$

Hence there exists a strategy to achieve 2ϵ advantage from the given parity attacker of advantage ϵ. If semantic security does not mean protection against such parity attacks, we can form a simple contradiction.

Hence, a semantically secure scheme is also secure against parity prediction.

2 Randomized Algorithms Warmup

1. (2 points) Start with rolling an n sided die. If it lands on $r > 1$, roll an r-sided die and repeat. If it lands on $r = 1$, then halt. What is the expected number of rolls before halting?

Solution: Consider a n sided die that is rolled once. With a probability of 1/n, it lands on 1 and the sequence would be terminated. In the cases that it does not land on 1, it lands on {2,3,4,.....,n} with the probability distribution of: (1/n,1/n,1/n,.....,1/n).

If the die lands on one of the numbers, say i, from {2,3,4,.....,n} then the expected number of further rolls to reach 1 would be same as the expected number of rolls when we start with a i-sided die and follow the same rules.

If we define D_n as the random variable denoting the expected number of such rolls to land on 1, we have:

$$E(D_n) = \frac{1}{n} * (1) + \frac{1}{n} * (1 + E(D_2)) + \frac{1}{n} * (1 + E(D_3)) + \frac{1}{n} * (1 + E(D_4)) + ... + \frac{1}{n} * (1 + E(D_n))$$

Similarly, we have:

$$E(D_{n-1}) = \frac{1}{n-1} * (1) + \frac{1}{n-1} * (1 + E(D_2)) + \frac{1}{n-1} * (1 + E(D_3)) + ... + \frac{1}{n-1} * (1 + E(D_{n-1}))$$

Multiplying the former equation with 'n', the latter equation with 'n-1' and subtracting gives:

$$n * E(D_n) - (n-1) * E(D_{n-1}) = 1 + E(D_n)$$

$$E(D_n) - E(D_{n-1}) = \frac{1}{n-1}$$

From here, it is easy to show that E(Dn) = $\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n-1}$

2. (2 points) Prove the following "success amplification" theorem for Monte Carlo randomized algorithms (one whose output might be incorrect but will always halt in polynomial time).

Let P and Q be polynomials and A be a probabilistic algorithm computing a function f such that: $\Pr[A(x) = f(x)] > \frac{1}{2} + \frac{1}{P(|x|)}$. *Then, by repeating the computation, prove that you can amplify the success probability to* $\Pr[A(x) = f(x)] > 1 - \frac{1}{Q(|x|)}$. *How many repetitions does it take?*

Solution: Let E be the event that A(x) = F(x).
We have Pr(E) > 1/2 + 1/P($|x|$).

Now let us say that we repeat the experiment n times and consider the majority of results. We will show that this majority must be more the real F(x) with a probability of more than $1 - 1/Q(|x|)$.
Consider a general term in the binomial expansion upto half (corresponding to event: majority of results correspond to the actual result).

A generic term in this decomposition can be described by the inequality: $T_i =\leq \binom{k}{i}(1/4 - 1/P(|x|))^{k/2}$

Considering the summation upto k/2 terms: we get

$Pr(A(x) = f(x)) > 1 - \frac{P(|x|)^2}{4k}$
From here, if we set $k > 1/4 * P(|x|)^2 * Q(|x|)$, we have total probability of actualy F(x) coming in the majority as $1 - \frac{1}{Q(|x|)}$. Thus $k > 1/4 * P(|x|)^2 * Q(|x|)$. Hence we have shown a strategy for success amplification.