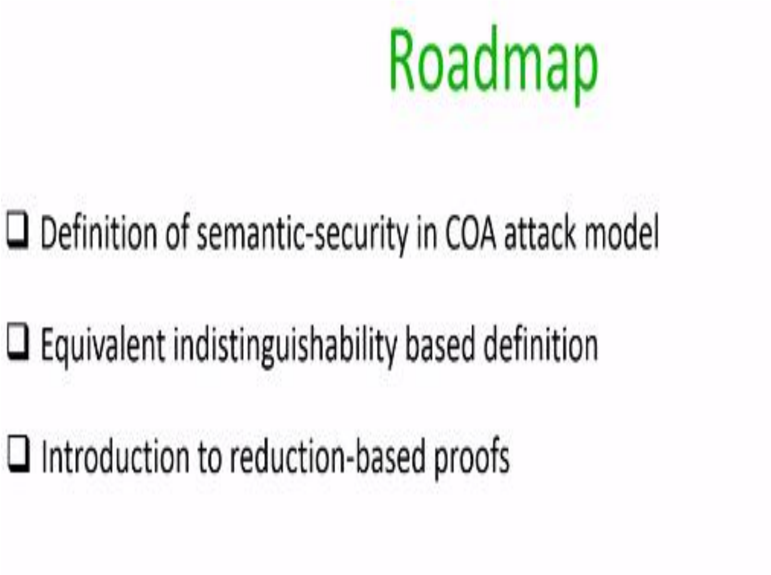
Foundations of Cryptography

Prof. Dr. Ashish Choudhury
(Former) Infosys Foundation Career Development Chair Professor
International Institute of Information Technology – Bangalore

Lecture – 7
Semantic Security

(Refer Slide Time: 00:32)



Roadmap

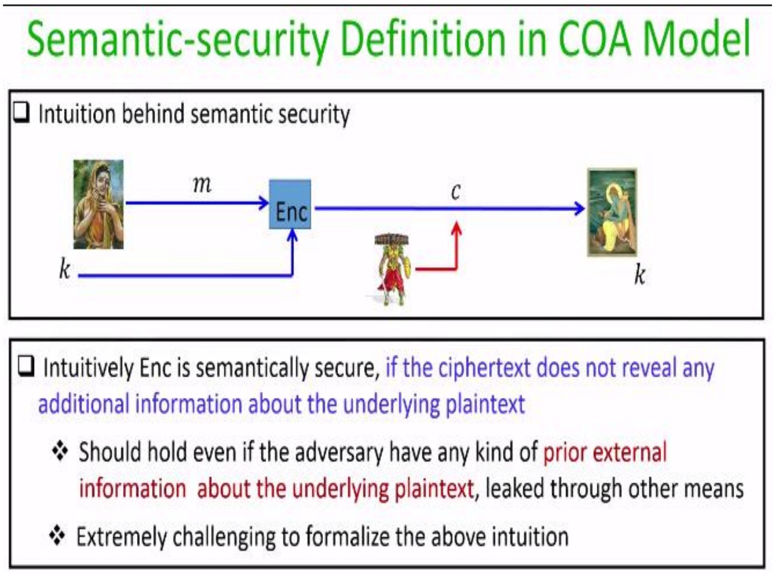
❑ Definition of semantic-security in COA attack model

❑ Equivalent indistinguishability based definition

❑ Introduction to reduction-based proofs

Hello, everyone, welcome to lecture 7. The plan for this lecture is as follows. We will define the notion of semantic security in the ciphertext only attack model and we will see an equivalent version of this definition based on indistinguishability game and we will also introduce reduction-based proofs which is central to cryptography.

(Refer Slide Time: 00:51)



Semantic-security Definition in COA Model

❑ Intuition behind semantic security

❑ Intuitively Enc is semantically secure, if the ciphertext does not reveal any additional information about the underlying plaintext

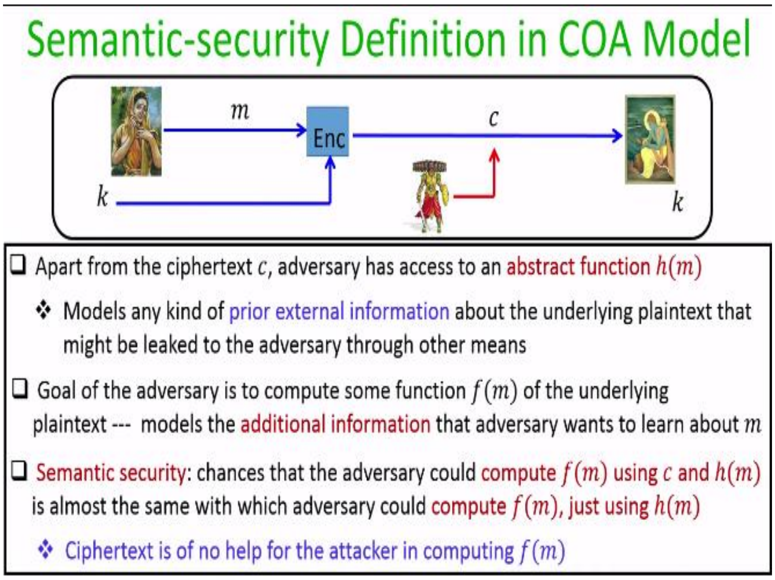
- ❖ Should hold even if the adversary have any kind of prior external information about the underlying plaintext, leaked through other means
- ❖ Extremely challenging to formalize the above intuition

So, let us begin with the semantic-security definition in the ciphertext only attack model and the scenario is the following. So, we are in the ciphertext only attack model where we have an adversary, and now we will consider a computationally bounded adversary. Because remember, in the last lecture we have discussed that if key reusability is your ultimate goal, then you have to ensure that your adversary is computationally bound.

So, we assume we have a computationally bounded adversary, who is seeing a ciphertext c of some unknown message m encrypted by the sender using an unknown key k as per the encryption algorithm, where the steps of the encryption algorithm is known to the adversary. Intuitively, we will say that our encryption process is semantically secure in this ciphertext only attack model if the ciphertext does not reveal any additional information about the underlying plaintext to the attacker.

Moreover, this should hold even if the adversary have any kind of prior external information about the underlying plaintext, which could have been leaked to the attacker by any other means before the ciphertext have been communicated. So even though this intuition is very straightforward to understand, it is extremely challenging to formalize the above intuition. So let us proceed to formalize this intuition, right.

(Refer Slide Time: 02:11)



Semantic-security Definition in COA Model

❑ Apart from the ciphertext c, adversary has access to an abstract function $h(m)$

❖ Models any kind of prior external information about the underlying plaintext that might be leaked to the adversary through other means

❑ Goal of the adversary is to compute some function $f(m)$ of the underlying plaintext --- models the additional information that adversary wants to learn about m

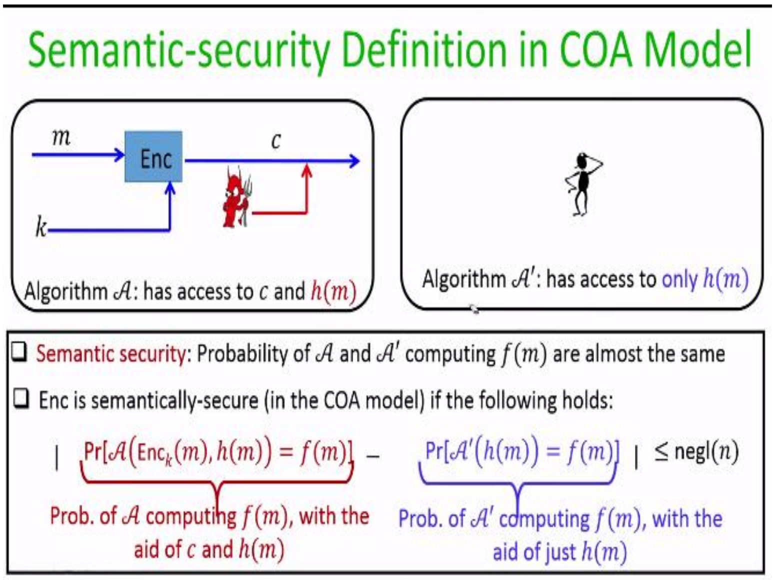
❑ Semantic security: chances that the adversary could compute $f(m)$ using c and $h(m)$ is almost the same with which adversary could compute $f(m)$, just using $h(m)$

❖ Ciphertext is of no help for the attacker in computing $f(m)$

So, we first introduce an abstract function here, namely h(m), which models any kind of prior external information about the underlying plaintext, which might be leaked to the adversary through other means before any ciphertext have been communicated. So this h(m) is kind of some history function. There might be in some context or might be a scenario where adversary might have absolutely no prior external information, in that case, my function h(m) will be an empty function, but there might be a scenario where adversary might have some prior external information about the underlying plaintext, which has been leaked to the adversary through some other means. So whatever is the case, we introduce this abstract function to model this prior external information about the underlying plaintext, which the adversary has.

We next introduce another function f(m), which basically models the additional information about the underlying plaintext, which adversary would like to compute after seeing the ciphertext or which adversary would like to know, right? So this models the additional information and intuitively, the goal of semantic security is to ensure the following. We say that our encryption process is semantically secure if the probability with which adversary could compute this additional information, namely the function f(m) by using the ciphertext c and using the help of the history function or the prior information is almost the same with which the adversary could have computed the function f(m) by just using the prior information in the absence of the ciphertext. If that is the case, then the implication that we get here is that ciphertext is of no help whatsoever for the attacker in computing f(m).

(Refer Slide Time: 04:04)



Semantic-security Definition in COA Model

Algorithm $\mathcal{A}$: has access to c and $h(m)$

Algorithm $\mathcal{A}'$: has access to only $h(m)$

- ❑ Semantic security: Probability of $\mathcal{A}$ and $\mathcal{A}'$ computing $f(m)$ are almost the same
- ❑ Enc is semantically-secure (in the COA model) if the following holds:

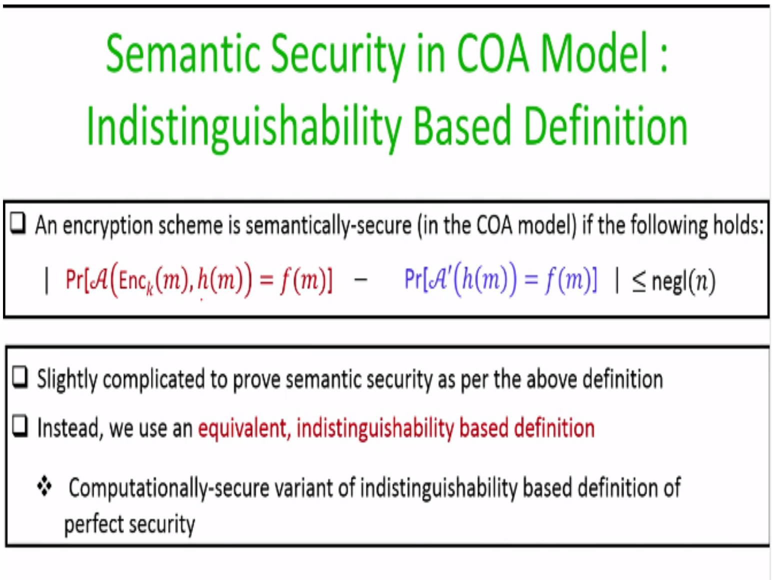
$$\left| \underbrace{\Pr[\mathcal{A}(\mathsf{Enc}_k(m), h(m)) = f(m)]}_{\text{Prob. of } \mathcal{A} \text{ computing } f(m)\text{, with the aid of } c \text{ and } h(m)} - \underbrace{\Pr[\mathcal{A}'(h(m)) = f(m)]}_{\text{Prob. of } \mathcal{A}' \text{ computing } f(m)\text{, with the aid of just } h(m)} \right| \leq \mathsf{negl}(n)$$

What I mean by this is pictorially the following. So you imagine that in this world, we have an adversary, who is actually seeing a ciphertext, which is an encryption of some unknown message under the unknown key and the adversary also have access to the history function, namely any kind of prior information, which would have been leaked to the adversary through some external mechanism without the knowledge of the ciphertext. So the adversary in this world is called A.

You compare by imagining another world where we have again another adversary, say A', who does not see the ciphertext and this adversary A' has access only to the history function, namely the prior information about the underlying plaintext, which sender might communicate over the channel. Now, the intuition behind semantic security is that the probability with which A and A' could compute the f(m), namely the additional information about the underlying plaintext are almost the same, namely, we will say that our encryption process is semantically secure in the ciphertext attack model if the absolute difference between the following two probabilities is upper bounded by some negligible function. So let us see closer. Let us have a closer look into this respective probabilities, namely :
$Pr[\mathcal{A}\left(Enc_k(m), h(m)\right) = f(m)]$ and $Pr[\mathcal{A}'(h(m)) = f(m)]$
So your first probability is the probability with which the adversary A, namely the adversary in the first world, outputs the value of f(m), where the adversary is given the ciphertext as well as the history function. Whereas the second probability is the probability with which the adversary in the second world, namely the adversary A', computes the value of f(m) just using the value of history function. So if the absolute difference between these two probabilities is upper bounded by a negligible probability, then what it means is that whatever adversary could have computed by seeing the ciphertext, namely whatever the adversary could have known about f(m) using the help of c with almost the same probability adversary could have computed f(m) without actually seeing the c.

If that is the case, then it means that our encryption process is so good that ciphertext is kind of some random bit strings, and it helps or provides no kind of aid to the adversary in computing f(m) with a significant advantage, that is what is the intuition behind the notion of semantic security, right.

(Refer Slide Time: 06:42)



Semantic Security in COA Model : Indistinguishability Based Definition

❑ An encryption scheme is semantically-secure (in the COA model) if the following holds:

$$\left| \Pr[\mathcal{A}(\mathsf{Enc}_k(m), h(m)) = f(m)] - \Pr[\mathcal{A}'(h(m)) = f(m)] \right| \leq \mathsf{negl}(n)$$

❑ Slightly complicated to prove semantic security as per the above definition

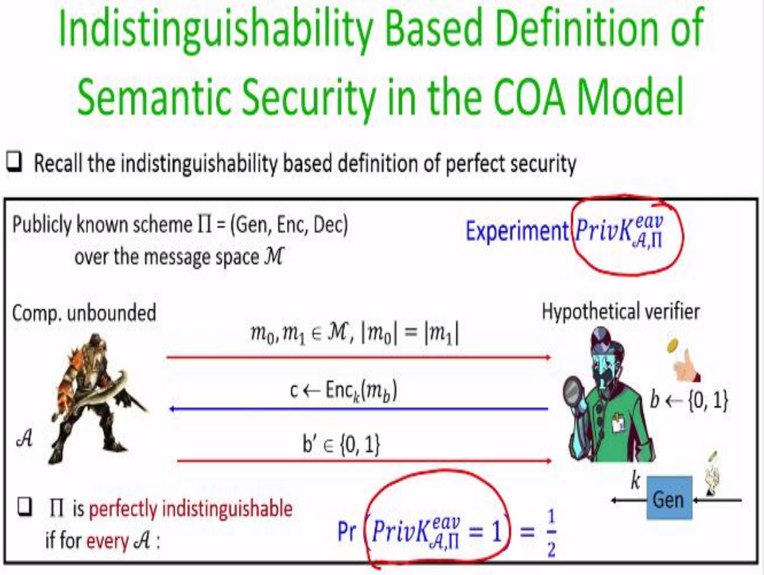
❑ Instead, we use an equivalent, indistinguishability based definition

❖ Computationally-secure variant of indistinguishability based definition of perfect security

So this is the original definition of semantic security, and it turns out that if we want to prove the semantic security of an arbitrary encryption process as per this original definition, then this is slightly complicated, where because here you have to bring history function as well as here you have to bring the arbitrary function f(m) which adversary would like to compute. Instead, what we are going to do is we will see an equivalent version of this definition based on indistinguishability based experiment.

This alternate definition based on indistinguishability based experiment, you can imagine that it is the computationally secure variant of indistinguishability based definition of perfect secrecy, right.

(Refer Slide Time: 07:27)



Indistinguishability Based Definition of Semantic Security in the COA Model

❑ Recall the indistinguishability based definition of perfect security

Publicly known scheme Π = (Gen, Enc, Dec) over the message space $\mathcal{M}$

Experiment $PrivK^{eav}_{\mathcal{A},\Pi}$

❑ Π is perfectly indistinguishable if for every $\mathcal{A}$:

$$\Pr\left(PrivK^{eav}_{\mathcal{A},\Pi} = 1\right) = \frac{1}{2}$$

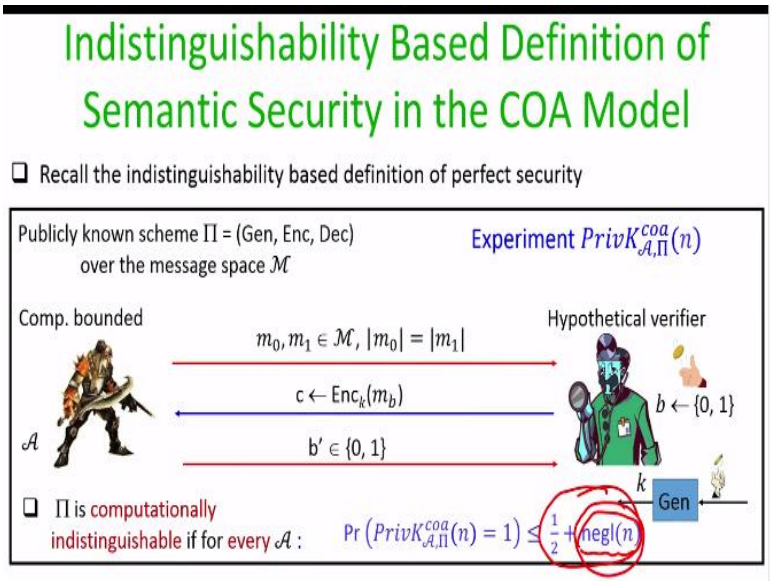
So, let us first recall the indistinguishability based definition that we use to define perfect secrecy. So, the essence of that indistinguishability based definition for defining perfect secrecy is that if you have a scenario where a sender has 2 messages m_0 or m_1 and it has randomly encrypted one of those messages, and if adversary is aware of the fact that the sender has either encrypted m_0 or m_1, then even after having this prior information and seeing the ciphertext c, adversary should not be able to identify what has been encrypted in the ciphertext c with probability better than half.

That was the intuition that we wanted to capture through the indistinguishability based definition of perfect secrecy. This was captured very nicely by the following experiment, right. So this is the experiment, which we use to define the notion of perfect secrecy, where we have a publicly known scheme, namely a triplet of algorithms over some message space, and in the model of perfect secrecy, we had a computationally unbounded adversary, and the name of the experiment was this : $PrivK_{\mathcal{A},\Pi}^{eav}$. So just to recall the nomenclature of the experiment is as follows. PrivK denotes that we want to model an experiment, which in the context of a private key or symmetric encryption, eav means we are considering an adversary who is an eavesdropper, A is the name of the adversarial algorithm and Π is the name of the scheme, and in this experiment, the rules are as follows. Adversary is allowed to submit any pair of messages from the plaintext space with the restriction that the size of the two plaintext should be same, and the experiment or the verifier, the hypothetical verifier does the following:

It randomly generates a key by running a key generation algorithm and it randomly encrypts one of the messages using the key, and the challenge for the adversary is to identify what plaintext has been encrypted in the challenge ciphertext c, whether it is m_0 or m_1. So adversary outputs a bit, namely its guess about what exactly has been encrypted in the challenge ciphertext. We define the scheme Π to be perfectly secure or we say that a scheme is perfectly indistinguishable if the probability with which adversary could successfully identify what message has been encrypted is upper bounded by half.

So if adversary could successfully identify what message has been encrypted in the challenge ciphertext, then we say that adversary has won the game or we say that output of the experiment is equal to 1. That means this notation that output of the experiment is 1 denotes the probability that b' equal to b. So in the context of perfect secrecy, our requirement was that a probability adversary could identify b, b' equal to b correctly should be upper bounded by half.

(Refer Slide Time: 10:25)



Indistinguishability Based Definition of Semantic Security in the COA Model

❑ Recall the indistinguishability based definition of perfect security

Publicly known scheme Π = (Gen, Enc, Dec) over the message space $\mathcal{M}$

Experiment $PrivK^{coa}_{\mathcal{A},\Pi}(n)$

Comp. bounded

$\mathcal{A}$

Hypothetical verifier

$m_0, m_1 \in \mathcal{M}$, $|m_0| = |m_1|$

$c \leftarrow \mathsf{Enc}_k(m_b)$

$b' \in \{0, 1\}$

$b \leftarrow \{0, 1\}$

k

Gen

❑ Π is computationally indistinguishable if for every $\mathcal{A}$: $\Pr\left(PrivK^{coa}_{\mathcal{A},\Pi}(n) = 1\right) \leq \frac{1}{2} + \mathsf{negl}(n)$

Now, let us see the indistinguishability based definition to model the notion of semantic security in the ciphertext only attack model. We will make the following changes. The first change that we are going to make is the following. Instead of assuming that our adversary is computationally unbounded, we will assume that our adversary is computationally bounded. This is to model the first relaxation that we agreed that we should make in computational security model right.

So remember the last lecture we discussed that if key reusability is your ultimate goal, then we should target to achieve security only against a computationally bounded adversary, namely an adversary whose running time is upper bounded by some polynomial function of the security parameter. So, that is why we made this first change in the experiment, we will not be considering an adversary whose running time is computationally unbounded. Consequently, the name of the experiment is going to be $PrivK^{coa}_{\mathcal{A},\Pi}\ (n)$.

So, instead of saying EAV, I am now calling this experiment COA to denote that this is the indistinguishability experiment in the ciphertext only attack model and the second difference here in the nomenclature is that I am now introducing this security parameter n because the running time of the adversary is going to be upper bounded by a polynomial function of the security parameter, whereas if you see in the nomenclature for the indistinguishability based experiment in the world of perfect secrecy, no such security parameter was there because our adversary was allowed to have unlimited running time.

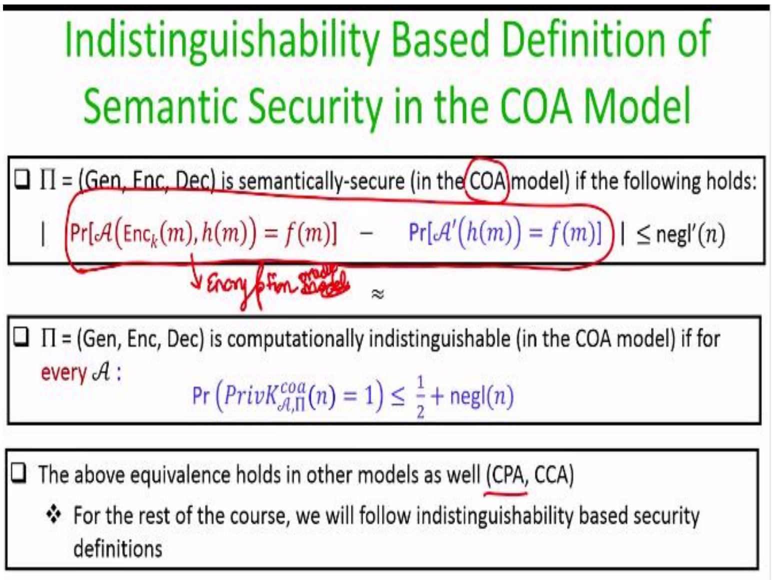
So this is the second relaxation. This is the second change in the experiment, and the third change is that instead of saying that our encryption process is perfectly indistinguishable, we will say that our encryption process is computationally indistinguishable. If the probability with which adversary could correctly identify what has been encrypted in the challenge ciphertext is upper bounded by some negligible function plus half, that is a third change we are going to make in our experiment, right.

So in the experiment for perfect secrecy, the requirement was that adversary should not be able to identify what has been encrypted in the challenge ciphertext with probability better than half, but now we are giving the adversary extra negligible advantage to correctly identify what has been encrypted in the challenge ciphertext c. This extra advantage, namely an advantage of negligible function and advantage of some negligible probability is to model the second relaxation that we have to make in the model of computational security if the key reusability is your ultimate goal.

So again recall in the last lecture, we have seen that if you want to design a scheme where key reusability is your ultimate goal, that instead of demanding that adversary should not learn anything additional, you should be willing to let the adversary learn something about your underlying message or to let the adversary break your scheme with some additional probability and that additional probability should be so small in which it should be a negligible probability, which for most practical purposes you can ignore it off.

So, that is why I am bringing this additional advantage of negligible function of n in my security definition. So, that is the computationally secure indistinguishability based version experiment in the ciphertext only attack model. So the essence of this experiment is the following. What we want to capture through this experiment is the following. If you have a scenario where a sender is having a pair of message, say m_0 and m_1 and if the adversary is aware of this where our adversary is computationally bounded, if one of these 2 messages m_0 or m_1 has been encrypted by the sender and communicated to the receiver and our adversary intercepts a ciphertext, then we require the following property from our encryption process. We require that a computationally bounded adversary should not be able to identify whether the ciphertext c which he is seeing is an encryption of m_0 or whether it is an encryption of m_1 with probability better than half plus negligible. That is what is the scenario or real world scenario we are trying to capture through this experiment.

(Refer Slide Time: 14:33)



Indistinguishability Based Definition of Semantic Security in the COA Model

❑ Π = (Gen, Enc, Dec) is semantically-secure (in the COA model) if the following holds:

$$\left| \Pr[\mathcal{A}(\mathsf{Enc}_k(m), h(m)) = f(m)] \;-\; \Pr[\mathcal{A}'(h(m)) = f(m)] \right| \leq \mathsf{negl}'(n)$$

Encryption [illegible]

$\approx$

❑ Π = (Gen, Enc, Dec) is computationally indistinguishable (in the COA model) if for every $\mathcal{A}$:

$$\Pr\left(PrivK^{coa}_{\mathcal{A},\Pi}(n) = 1\right) \leq \frac{1}{2} + \mathsf{negl}(n)$$

❑ The above equivalence holds in other models as well (CPA, CCA)

❖ For the rest of the course, we will follow indistinguishability based security definitions

So now we have the following 2 definitions. The first definition is actually the original definition of semantic security in the ciphertext only attack model, where we want to capture that the advantage of the adversary in first world and the adversary in the second world is upper bounded by negligible probability, whereas the second definition is the indistinguishability based definition. It turns out that both these two definitions are equivalent.

Namely if we have an encryption process which satisfies the first condition, then we can prove that for the same encryption process in the second condition also hold and vice versa. Namely, if we have an encryption process where the second condition hold, then for the same encryption process, the first condition also holds. I would like to stress the following. In the experiment, which we have discussed when I say that the probability that adversary correctly identifies what has been encrypted in the challenge ciphertext should be upper bounded by half plus negligible, then this probability is over the randomness of the experiment. So remember that the experiment could choose the message m_0 to be encrypted in the challenge ciphertext with probability half and with probability half the experiment or the verifier could choose the message m_1 to be encrypted in the ciphertext c. This probability of correctly identifying whatever has been encrypted in c should be also over the randomness of the adversary, right, because the entire experiment is going to be a randomized experiment, right.

So, as I said that these 2 notions of security or these 2 definitions are equivalent, and the proof that these 2 definitions are equivalent is slightly complicated and due to interest of time, we will not be going into the details of the proof. However, if you want to have a very high level overview of the proof, you can refer to the book by Katz and Lindell. Interestingly, it turns out that the equivalence of these 2 definitions holds in the other models as well, right

So, currently what we are considering is the ciphertext only attack model and in the ciphertext only attack model, the adversary has got access to the encryption of some message, but if I go to the higher attack model, by higher attack model means more powerful attack model, say the CPA attack model where apart from the ciphertext, adversary also gets access to the encryption oracle then we can have a corresponding semantically secure version of the definition that we are currently giving here for the COA model, right.

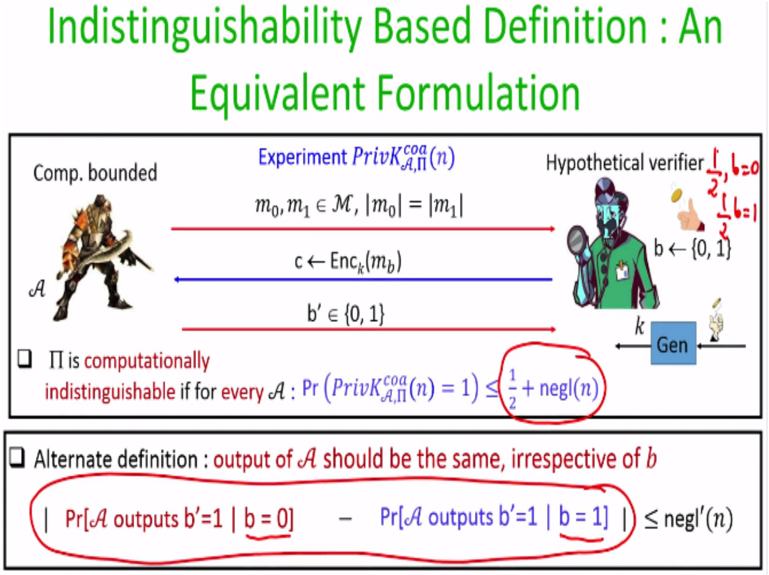
Namely, we would like to state that adversarial advantage or the difference of the absolute probabilities of adversary computing the function f(n) in the 2 worlds should be upper bounded by a negligible function, where in the first world apart from the ciphertext, adversary will also get access to the encryption oracle if we take this definition to the CPA attack model. In the same way if we take this definition to the CCA attack model, then apart from the ciphertext, adversary in the first world will have access to the encryption oracle, it will have access to the decryption oracle and so on.

So, we can come up with a semantically secure, we can come up with a version of the semantic security in the CPA model, in the CCA model and so on where the essence of the definition will be that the absolute difference between the two probabilities of adversity computing f(m) in the first world and f(m) in the second world should be upper bounded by a negligible probability, that will be the essence of the semantic security definition in CPA model, CCA model and so on.

It turns out that irrespective of the model, we can come up with a corresponding indistinguishability based definition and we can prove that the semantically secure version of the definition, the original version of the semantic security will be equivalent to the indistinguishability based definition. So, that is why for the rest of the course, we will not be seeing the original version of the semantic security definition.

We will be rather following the indistinguishability based security definition and depending upon whether we are in the CPA world, CCA world, we will enhance the indistinguishability based experiment, right.

(Refer Slide Time: 18:53)



Indistinguishability Based Definition : An Equivalent Formulation

Comp. bounded
$\mathcal{A}$
Experiment $PrivK^{coa}_{\mathcal{A},\Pi}(n)$
Hypothetical verifier $\frac{1}{2}, b=0$
$\frac{1}{2}, b=1$
$m_0, m_1 \in \mathcal{M}, |m_0| = |m_1|$
$c \leftarrow Enc_k(m_b)$
$b' \in \{0, 1\}$
$b \leftarrow \{0, 1\}$
k
Gen

❑ Π is computationally indistinguishable if for every $\mathcal{A}$: $\Pr\left(PrivK^{coa}_{\mathcal{A},\Pi}(n) = 1\right) \leq \frac{1}{2} + \text{negl}(n)$

❑ Alternate definition : output of $\mathcal{A}$ should be the same, irrespective of b

$$\left| \Pr[\mathcal{A} \text{ outputs } b'=1 \mid \underline{b = 0}] - \Pr[\mathcal{A} \text{ outputs } b'=1 \mid \underline{b = 1}] \right| \leq \text{negl}'(n)$$

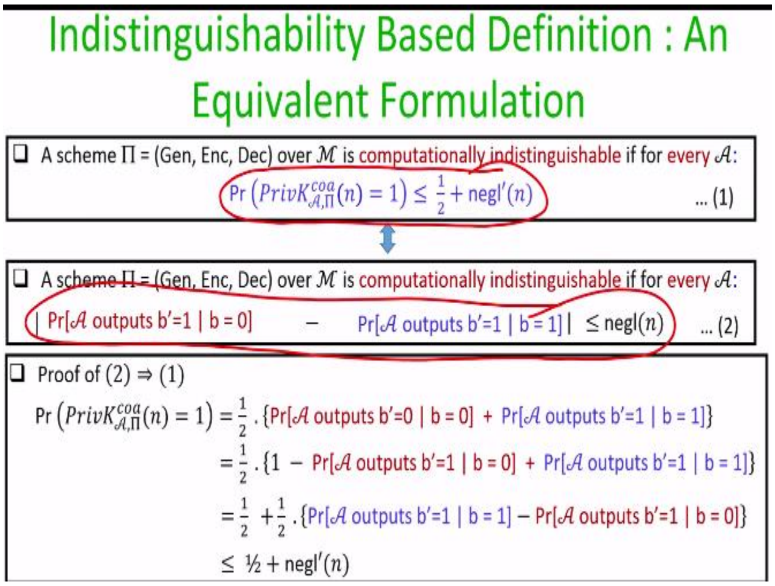
So, this is the indistinguishability based definition in the ciphertext only attack model and it turns out that we could come up with an alternate version of this definition, right. So the original definition requires you that the probability that our adversary is correctly able to identify the message that is encrypted in c should be upper bounded by half plus negligible, that is what is the original definition. The alternate definition demands that the output of the adversary should be same irrespective of what exactly is the message which has been encrypted in the challenge ciphertext c.

So, remember that since this indistinguishability based definition is a randomized experiment with probability half, my b could be equal to zero and with probability half the message which has been encrypted and ciphertext c could be m_1, right. The goal of the adversary is to identify whether it is m_0 which is encrypted in c or whether it is m_1 which has been encrypted in c. So this alternate definition demands that output of the adversary should be same irrespective of whether it is m_0 which is encrypted in c or whether it is m_1 which is encrypted in c.

More formally, it does not matter whether the message m_1 has been encrypted or message m_0 has been encrypted in c, in both the cases, adversary's output should be almost same except with a negligible probability. That means absolute advantage of the adversary distinguishing apart whether he is seeing an encryption of m_0 in ciphertext c or whether he is seeing an encryption of m_1 in the ciphertext c should be upper bounded by some negligible function.

If this is the case, then we say that our encryption process has indistinguishable encryption in the ciphertext only attack model, right. So, another interpretation of this difference of these **two** probabilities is upper bounded by a negligible probability is that a distinguishing advantage, you can view the difference between these 2 probabilities as the distinguishing advantage of our adversary, right? So the essence of this alternate definition is that the distinguishing advantage of our adversary in this experiment should be upper bounded by a negligible probability.

(Refer Slide Time: 21:12)



Indistinguishability Based Definition : An Equivalent Formulation

❑ A scheme Π = (Gen, Enc, Dec) over $\mathcal{M}$ is computationally indistinguishable if for every $\mathcal{A}$:

$$\Pr\left(PrivK^{coa}_{\mathcal{A},\Pi}(n) = 1\right) \leq \frac{1}{2} + \text{negl}'(n) \quad \text{... (1)}$$

❑ A scheme Π = (Gen, Enc, Dec) over $\mathcal{M}$ is computationally indistinguishable if for every $\mathcal{A}$:

$$\left| \Pr[\mathcal{A} \text{ outputs } b'=1 \mid b=0] - \Pr[\mathcal{A} \text{ outputs } b'=1 \mid b=1] \right| \leq \text{negl}(n) \quad \text{... (2)}$$

❑ Proof of (2) ⇒ (1)

$$\begin{aligned}\Pr\left(PrivK^{coa}_{\mathcal{A},\Pi}(n) = 1\right) &= \frac{1}{2}.\{\Pr[\mathcal{A} \text{ outputs } b'=0 \mid b=0] + \Pr[\mathcal{A} \text{ outputs } b'=1 \mid b=1]\} \\ &= \frac{1}{2}.\{1 - \Pr[\mathcal{A} \text{ outputs } b'=1 \mid b=0] + \Pr[\mathcal{A} \text{ outputs } b'=1 \mid b=1]\} \\ &= \frac{1}{2} + \frac{1}{2}.\{\Pr[\mathcal{A} \text{ outputs } b'=1 \mid b=1] - \Pr[\mathcal{A} \text{ outputs } b'=1 \mid b=0]\} \\ &\leq ½ + \text{negl}'(n)\end{aligned}$$

It turns out that these 2 versions of the indistinguishability based definitions are equivalent. Namely, we can say that our encryption process is computationally indistinguishable if the probability with which adversary could correctly output b equal to b**'** is upper bounded by half plus some negligible function. The second definition says that the distinguishing advantage of the attacker to distinguish apart whether he is seeing an encryption of m_0 or whether he is seeing an encryption of m_1 should be upper bounded by a negligible probability.

It turns out that both these 2 conditions are equivalent. Namely, we can prove that if we have an encryption process Π where the first condition holds and for the same encryption process, the second condition also holds and vice versa. So, what we are going to do is we will follow the implication in the direction that the condition 2 implies condition 1, namely we will assume that say we have an encryption process where condition 2 holds, namely the distinguishing advantage of the attacker is upper bounded by some negligible probability.

If that is the case, then we are going to show that irrespective of the way adversary participates in the indistinguishability based experiment, the probability that adversary could correctly identify b equal to b' or it ensures b equal to b' is upper bounded by half plus some negligible function. So, let us prove that. So, what is the probability that in the indistinguishability based experiment, adversary outputs b equal to b' because if adversary outputs b equal to b', that is what is the interpretation that the experiment outputs 1.

Now, if you recall, there are 2 versions of the experiment. One version of the experiment where the challenger or the experiment has selected message m_0 in the ciphertext c, that means m_0 has been encrypted in c and the second version of the experiment is when the message m_1 has been encrypted in c. It turns out that with probability half, the challenger could use m_0 to encrypt in c and the probability half it could use m_1 to encrypt in c.

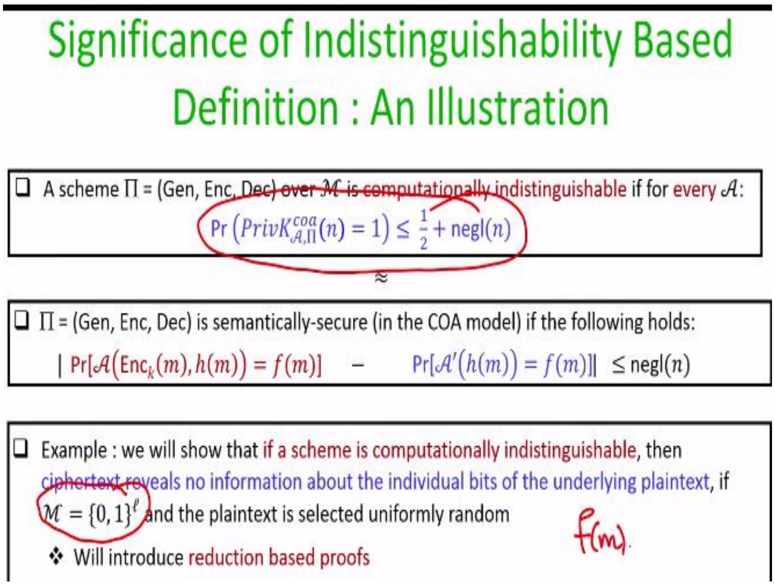
So overall, the probability that the experiment adversary outputs b equal to b' is he should output b equal to b' in the case when b = 0 and he should output b = b' even for the case when b = 1. Both these events b = 0 and b = 1 can occur with probability half. So that is why I am taking half as common, right. Now, what I am going to do here is the following. I am going to rewrite $Pr[\mathcal{A} \text{ outputs b}' = 0 \mid \text{b} = 0]$ as 1 minus the complimentary probability as, namely $\{1 - Pr[\mathcal{A} \text{ outputs b}' = 1 \mid \text{b} = 0]\}$, which means even though the message m_0 has been encrypted in the challenge ciphertext, and if I subtract that probability from 1, then I get the probability with which adversary could identify that indeed m_0 has been encrypted in c, given that indeed $m_{0,}$ has been encrypted in c, right. So, that is what is the substitution I have done. Now, what I can do is I can take this 1/2 inside the bracket, and as a result, by rearranging the term I get this given expression :

$$\frac{1}{2} + \frac{1}{2} . \{Pr[\mathcal{A} \text{ outputs b}' = 1 \mid \text{b} = 1] - Pr[\mathcal{A} \text{ outputs b}' = 1 \mid \text{b} = 0]\}$$

Now I make use of the fact that I am assuming that for my encryption process, the distinguishing advantage of the attacker is upper bounded by some negligible probability. So if you see the highlighted thing : $\{\boldsymbol{Pr}[\boldsymbol{\mathcal{A}} \textbf{ outputs b}' = \mathbf{1} \mid \mathbf{b} = \mathbf{1}] - \boldsymbol{Pr}[\boldsymbol{\mathcal{A}} \textbf{ outputs b}' = \mathbf{1} \mid \mathbf{b} = \mathbf{0}]\}$, this is nothing but the absolute difference between the 2 probabilities which highlights actually the distinguishing advantage of the attacker, namely, the probability with which it can distinguish apart whether it is seeing an encryption of m_0 or whether it is seeing an encryption of m_1. As part of our assumption, we are assuming that the distinguishing advantage of the attacker is upper bounded by some negligible function. So, I can substitute this highlighted thing by some negligible function. So, what I obtain here is that the probability that adversary outputs b = b' in the experiment is upper bounded by half plus some other negligible function because 1/2 into negligible function, I can always replace by another negligible function, negl'.

So, what I have shown here is if the condition 2 holds for my arbitrary encryption process, then even the condition one holds as well. In the same way, we can prove that if condition 1 holds for our encryption process, then for the same encryption process condition 2 holds as well, I leave that as an exercise for you. That means both these 2 versions of the definitions are same, we can prove that we can either demand that condition 1 holds, if we want to say that our encryption process is COA secure as per the indistinguishability game or we can demand that the distinguishing advantage of the adversary should be upper bounded by a negligible probability. Both of them are equivalent. Throughout the course, depending upon the convenience, we can use any of these 2 versions of the indistinguishability based definition of the ciphertext only attack, right.

(Refer Slide Time: 26:20)



Significance of Indistinguishability Based Definition : An Illustration

❑ A scheme Π = (Gen, Enc, Dec) over $\mathcal{M}$ is computationally indistinguishable if for every $\mathcal{A}$:

$$\Pr\left(PrivK^{coa}_{\mathcal{A},\Pi}(n) = 1\right) \le \frac{1}{2} + \mathsf{negl}(n)$$

$\approx$

❑ Π = (Gen, Enc, Dec) is semantically-secure (in the COA model) if the following holds:

$$| \Pr[\mathcal{A}(\mathsf{Enc}_k(m), h(m)) = f(m)] \quad - \quad \Pr[\mathcal{A}'(h(m)) = f(m)]| \ \le \mathsf{negl}(n)$$

❑ Example : we will show that if a scheme is computationally indistinguishable, then ciphertext reveals no information about the individual bits of the underlying plaintext, if $\mathcal{M} = \{0,1\}^\ell$ and the plaintext is selected uniformly random

f(m).

❖ Will introduce reduction based proofs

So, here is the summary. So we have the original definition of semantic security, right. So, this is your definition of semantic security in the second equation in the slide above, and we have an indistinguishability based definition in the first equation, and even though we have not proved it formally, you have to believe me that both these 2 definitions, both these 2 conditions are equivalent to each other. So, now, you might be wondering that how can these 2 conditions definitions might be equivalent because in the original definition of semantic security, you have an abstract function. Namely the history function and you have some function f(m) which adversary would like to compute, whereas in the computationally indistinguishability based definition there is no such history function, no such function f(m) which adversary would like to compute. So, you might be wondering how come these 2 conditions are equivalent? So, what we are going to do next is I am going to take an illustration and we will consider the case where our encryption process satisfies the indistinguishability based definition.

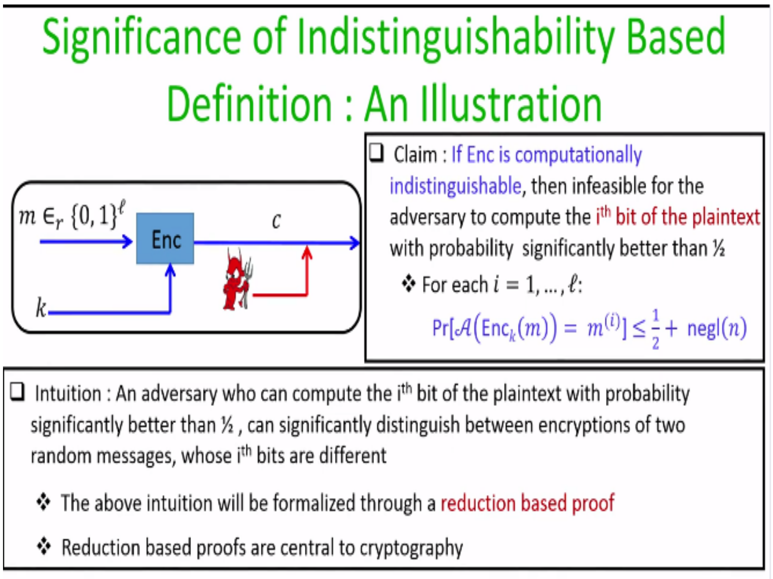
Namely, we will assume that our encryption process satisfies the condition $\Pr(PrivK_{\mathcal{A},\Pi}^{coa}\ (n) = 1) \leq \frac{1}{2} + \text{negl}(n)$, and we will assume that our distribution of the plaintext space is a uniform distribution over the set of messages, namely the set of l bit strings, and what we are going to show is that if our encryption process satisfies the indistinguishability based definition, then it implies that adversary by looking into the ciphertext cannot compute any bit of the underlying plaintext.

So, the function f(m) which adversary is interested to compute in this illustration is to compute the underlying bits of the message by looking into the ciphertext, and our goal will be to prove that if our encryption process satisfies the indistinguishability based definition, then the probability with which the adversary could compute this candidate f(m) is upper bounded by a negligible function, right. The proof strategy that we will use in this illustration is reduction based proof, which is kind of central to the cryptography because, in the rest of the course, almost all the proofs will be following this reduction based proof.

This reduction based proofs actually comes from another interesting branch of computer science namely complexity theory. Specifically, if you want to prove that a problem y is NP complete, then the way you proceed to prove that a problem y is NP complete is you show actually that any existing problem x in NP can be reduced to an instance of this new problem y and the implication of this reduction is that if you have a polynomial time solution for solving the problem y, then actually you have a polynomial time solution for the problem x as well, right?

So, that is what is the essence of this reduction. In the same way, we are going to use such kind of reduction based proofs in cryptography as well, where we will use an adversarial algorithm to break a scheme or to attack a scheme to solve or design another adversary or algorithm to break something else as well.

(Refer Slide Time: 29:36)



Significance of Indistinguishability Based Definition : An Illustration

❑ Claim : If Enc is computationally indistinguishable, then infeasible for the adversary to compute the i^th bit of the plaintext with probability significantly better than ½

❖ For each $i = 1, \ldots, \ell$:

$$\Pr[\mathcal{A}(\mathsf{Enc}_k(m)) = m^{(i)}] \leq \frac{1}{2} + \mathsf{negl}(n)$$

❑ Intuition : An adversary who can compute the i^th bit of the plaintext with probability significantly better than ½ , can significantly distinguish between encryptions of two random messages, whose i^th bits are different

❖ The above intuition will be formalized through a reduction based proof

❖ Reduction based proofs are central to cryptography

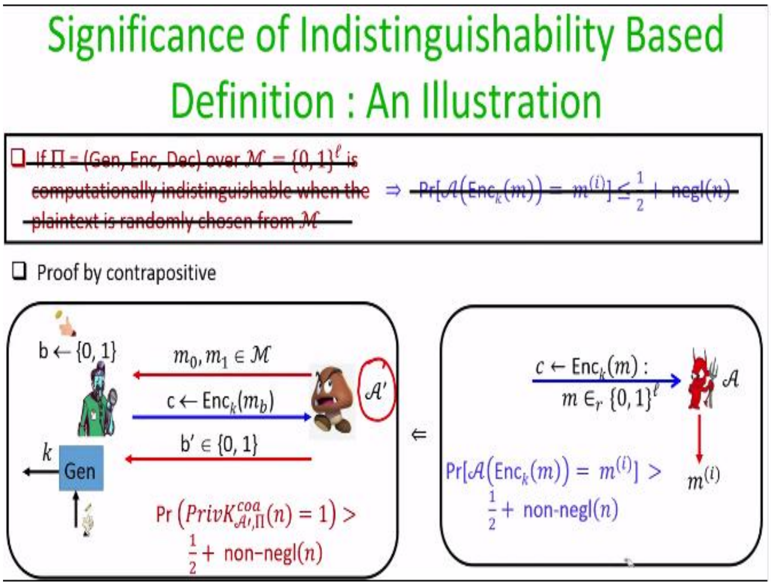
In more detail, what we are going to show in this illustration is the following, right. So imagine your underlying message, the message which sender has used for encryption is randomly selected from the set of *l* bit strings and we are in the ciphertext only attack model and our adversary has seen a ciphertext c. So, the claim that we are making here is that if your encryption process is computationally indistinguishable, that means it satisfies the notion of computational indistinguishability, then it is impossible for any polynomial time adversary to compute any of the underlying bits of the plaintext with probability significantly better than half. That is what we are going to show. That means for each i, i belonging to 1 to *l*, there is the probability that any algorithm A, which when given an encryption of an unknown message, which is randomly chosen from the set of *l* bit strings and encrypted using an unknown key that adversarial algorithm when given such a ciphertext outputs the ith bit of the underlying plaintext is upper bounded by half plus negligible.

So this m raised to power i actually denotes the ith bit of the underlying plaintext. So, this illustration basically is going to show you that if your encryption process actually satisfies the indistinguishability based definition, then indeed it is not possible for any polynomial time algorithm to compute any f(m) of his choice, where in this particular illustration, the function f(m) is computing the ith bit of the underlying plaintext. The basic intuition to prove this claim is the following.

Imagine that we have an adversary A, which when given an encryption of a random message could compute the ith bit of the underlying message with significant probability better than half. Then, using this algorithm or adversary, we can design another adversary, which can distinguish apart an encryption of message m_0 and an encryption of message m_1 where the ith bits are different with significant probability, but this will be a violation to the assumption that my encryption process is computationally indistinguishable.

Because when I say my encryption process is computationally indistinguishable, that means there exist no adversary, which can distinguish apart an encryption of message m_0 and an encryption of message m_1 whose ith bits are different. That is what is the basic intuition of the claim that we are making here. However, this claim has to be formalized by giving a reduction based proof and let us see the details of the reduction based proof.

(Refer Slide Time: 32:21)



Significance of Indistinguishability Based Definition : An Illustration

❑ ~~If Π = (Gen, Enc, Dec) over $\mathcal{M} = \{0,1\}^{\ell}$ is computationally indistinguishable when the plaintext is randomly chosen from $\mathcal{M}$~~ $\Rightarrow$ ~~$\Pr[\mathcal{A}(\mathsf{Enc}_k(m)) = m^{(i)}] \leq \frac{1}{2} + \mathsf{negl}(n)$~~

❑ Proof by contrapositive

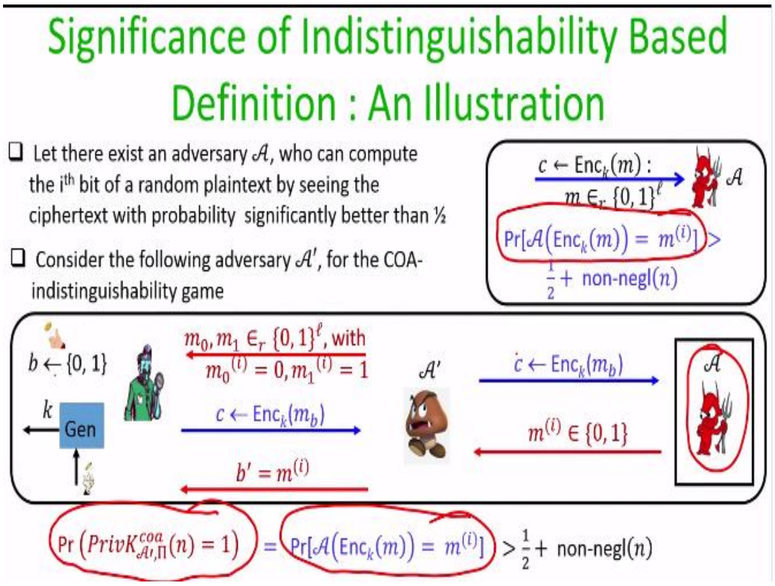
$b \leftarrow \{0,1\}$
$m_0, m_1 \in \mathcal{M}$
$c \leftarrow \mathsf{Enc}_k(m_b)$
$b' \in \{0,1\}$
k
Gen
$\mathcal{A}'$
$\Pr\left(PrivK^{coa}_{\mathcal{A}',\Pi}(n) = 1\right) > \frac{1}{2} + \text{non-negl}(n)$
$\Leftarrow$
$c \leftarrow \mathsf{Enc}_k(m)$: $m \in_r \{0,1\}^{\ell}$
$\mathcal{A}$
$m^{(i)}$
$\Pr[\mathcal{A}(\mathsf{Enc}_k(m)) = m^{(i)}] > \frac{1}{2} + \text{non-negl}(n)$

So the claim that we are making is formally stated here. So imagine what the claim states that if you have an encryption process over the message space of *l* bit strings, and if the underlying message is selected uniformly randomly and if your encryption process is computationally indistinguishable, then it implies there exist no adversary, which when given an encryption of a random unknown message under random key can compute the ith bit of the underlying plaintext with significant probability better than half.

The proof strategy here will be the proof by contrapositive, right. Namely, we will assume that suppose there exist an adversary, who can significantly compute the ith bit of the underlying plaintext where the underlying plaintext is randomly chosen from the set of *l* bit strings and underlying plaintext is not known to the adversary and also the key is not known to the adversary. That means, the property of this adversary is if it is given an encryption or ciphertext c of an unknown randomly chosen message m, it could correctly identify or it could correctly output the ith bit of the underlying plaintext with probability significantly better than half.

That is what we mean when I say that there exist an adversary who can correctly identify the ith bit of the underlying plaintext. Now assuming the existence of this adversary, what we are going to show is that the claim that we are making our encryption process, namely the claim that our encryption process is computationally indistinguishable, also does not hold. Namely, we will design an adversary A**'**, which can participate in an instance of the indistinguishability based game in the COA attack model and with significant probability better than half it can distinguish apart whether it is seeing an encryption of m_0 or whether it is seeing an encryption of m_1, right. So, that means, with whatever advantage A could identify the ith bit of the underlying plaintext, namely with whatever advantage our adversary A could compute the function f(m), so remember the function f(m) here in this case is computing the ith bit of the underlying plaintext. So, if there exist an adversary A who can significantly compute value of function f(m) with non-negligible probability better than half, then with almost the same advantage, we can design another algorithm A**'** who can win an instance of COA indistinguishability experiment. That is what we are going to show through this reduction. So let us see the details of the reduction here.

(Refer Slide Time: 35:08)



Significance of Indistinguishability Based Definition : An Illustration

❑ Let there exist an adversary $\mathcal{A}$, who can compute the ith bit of a random plaintext by seeing the ciphertext with probability significantly better than ½

❑ Consider the following adversary $\mathcal{A}'$, for the COA-indistinguishability game

$c \leftarrow \mathrm{Enc}_k(m)$:
$m \in_r \{0,1\}^{\ell}$
$\mathcal{A}$
$\Pr[\mathcal{A}(\mathrm{Enc}_k(m)) = m^{(i)}] > \frac{1}{2} + \text{non-negl}(n)$
$b \leftarrow \{0,1\}$
k
Gen
$m_0, m_1 \in_r \{0,1\}^{\ell}$, with $m_0^{(i)} = 0, m_1^{(i)} = 1$
$c \leftarrow \mathrm{Enc}_k(m_b)$
$b' = m^{(i)}$
$\mathcal{A}'$
$c \leftarrow \mathrm{Enc}_k(m_b)$
$m^{(i)} \in \{0,1\}$
$\mathcal{A}$

$$\Pr\left(PrivK^{coa}_{\mathcal{A}',\Pi}(n) = 1\right) = \Pr[\mathcal{A}(\mathrm{Enc}_k(m)) = m^{(i)}] > \frac{1}{2} + \text{non-negl}(n)$$

So assume we have an adversary A and the property of the adversary is that if it sees an encryption of a random message from the plaintext space of *l* bit strings, then it could identify the ith bit of the underlying plaintext with probability half plus non-negligible, where i ranges from 1 to *l*. So, for simplicity, you can imagine i is equal to 1, that means, the property of this adversary is if you give him an encryption of randomly chosen message, where the message is not known to him and underlying key is not known to him, then this adversary in polynomial time can correctly output the first bit of the message because I am assuming i equal to 1, this adversary could output the first bit of the message with probability half plus some non-negligible function. Now, assuming the existence of such an adversary A, we design adversary A' who participates in an instance of COA indistinguishability game as follows. So as per the steps of the COA indistinguishability game, the adversary A' has to submit two messages from the plaintext space.

What the adversary A' does is, it selects a pair of messages m$_0$ and m$_1$, and remember as per the steps of the indistinguishability game, the adversary A' is free to choose any pair of message from the plaintext space. So, the adversary A' here cleverly chooses the pair of messages here, the message m$_0$ as ith bit being 0 and the message m$_1$ as the ith bit being 1. That means, the message m$_0$ and m$_1$ differ in their ith bit, apart from that, all other bits of the messages m$_0$ and m$_1$ are randomly chosen.

That is what is the pair of messages which adversary submits to the experiment. Now, what the experiment does is it runs the key generation algorithm and it decides to either encrypt the message m$_0$ or to encrypt the message m$_1$. So, imagine the adversary A' has given the encryption of the message m_b where b is the index of the message which has been encrypted by the experiment, and the goal of the attacker A' is to identify whether it is m_0 which has been encrypted in c or whether it is m_1 which has been encrypted in c, right.

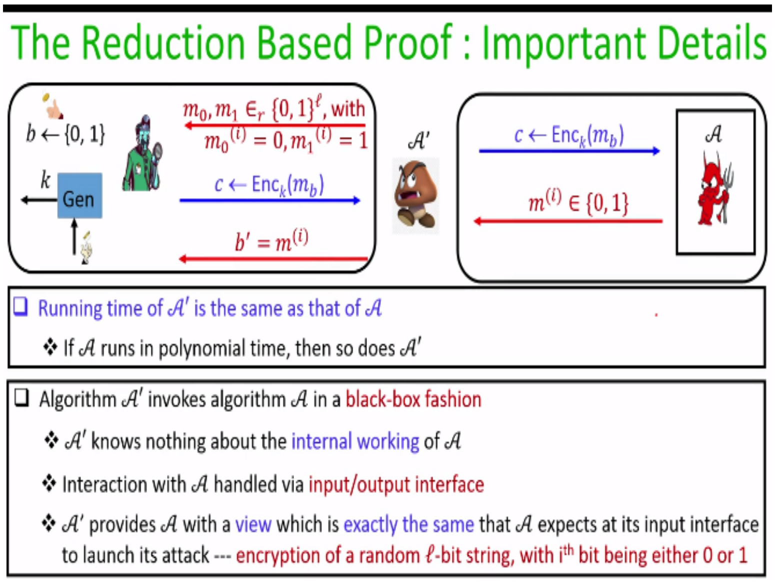
That is what is the goal of the attacker, whether it is an encryption of m_0 or whether it is an encryption of m_1. Now, what this attacker is going to do is it is going to take the help of the adversarial algorithm A, which we have assumed that it exists, right. Namely, A' invokes the algorithm A, namely it creates an instance of an encryption for the algorithm A by supplying a ciphertext c and where the ciphertext c is the same ciphertext which the adversary A' has obtained in the experiment, the COA experiment.

Namely the adversary A' here challenges the adversary A here to identify the i^{th} bit of the underlying message which has been encrypted in c. Now, what the algorithm A is going to do is it is going to identify the i^{th} bit of the underlying message which has been encrypted in the challenge ciphertext c, right? Because the property of the algorithm A is that if you give him an encryption of a random message, then it can tell you the i^{th} bit of the underlying message whether it is 0 or 1.

Now, depending upon what output the adversarial algorithm A' obtain from A, right, the adversarial algorithm $A^{'}$ comes back to the instance of the COA experiment, and it outputs b' = m_i. Namely, the response or the output from the adversarial algorithm A' in the COA experiment is the same bit which the adversarial algorithm A has output after seeing the ciphertext c, right. Now that my claim here is the probability that the algorithm A' outputs the correct message which has been encrypted in c, namely the probability with which A' outputs b = b' in the COA indistinguishability game is exactly the same with which the adversarial algorithm A could correctly output the i^{th} bit of the underlying message which has been encrypted in the ciphertext c. Mathematically, what I mean to say is the probability with which the adversary A' can win the COA experiment is exactly the same with which the probability with which the adversary A could identify the i^{th} bit of the underlying message which has been encrypted in the challenge ciphertext.

This is because the view of the adversary A here, which A' is providing to A is exactly the same as A expects to launch its attack, right. So, let me elaborate on this in the next slide.

(Refer Slide Time: 40:07)



The Reduction Based Proof : Important Details

$b \leftarrow \{0, 1\}$

k

Gen

$m_0, m_1 \in_r \{0,1\}^\ell$, with $m_0^{(i)} = 0, m_1^{(i)} = 1$

$c \leftarrow \mathsf{Enc}_k(m_b)$

$b' = m^{(i)}$

$\mathcal{A}'$

$c \leftarrow \mathsf{Enc}_k(m_b)$

$m^{(i)} \in \{0, 1\}$

$\mathcal{A}$

❑ Running time of $\mathcal{A}'$ is the same as that of $\mathcal{A}$

- ❖ If $\mathcal{A}$ runs in polynomial time, then so does $\mathcal{A}'$

❑ Algorithm $\mathcal{A}'$ invokes algorithm $\mathcal{A}$ in a black-box fashion

- ❖ $\mathcal{A}'$ knows nothing about the internal working of $\mathcal{A}$
- ❖ Interaction with $\mathcal{A}$ handled via input/output interface
- ❖ $\mathcal{A}'$ provides $\mathcal{A}$ with a view which is exactly the same that $\mathcal{A}$ expects at its input interface to launch its attack --- encryption of a random ℓ-bit string, with i^th bit being either 0 or 1

So, this is the reduction here. So, if you can see here this part of the experiment is the COA experiment. This is an instance of COA indistinguishability experiment, and in this part, A' is basically taking help of A, right. So, the first observation in this reduction is the following. The running time of our adversary algorithm A' is almost the same as the running time of the adversarial algorithm A because what is the running time of our algorithm A', it has to submit a pair of random messages which it can do in polynomial amount of time, and then it has to invoke the algorithm A.

So whatever is the running time of the algorithm A, if the running time of the algorithm A is polynomial, then the running time of A' invoking A is also going to be a polynomial time. Then finally A' outputs a bit b' which is going to take a polynomial amount of time. So, if at all the running time of my algorithm A, which I am assuming to exist is polynomial time, the running time of my algorithm A' is also polynomial time.

The second thing here is the algorithm A' invoke algorithm A in a black-box fashion. Namely, A' is not going to know that what exactly is the way what algorithm the algorithm A is following to identify the i^{th} bit of the underlying message which has been encrypted in c. So you can imagine there is some kind of interface here. There is an input interface through which A' can provide a challenge to A and there is an output interface through which A' provides its output. Namely, it says what exactly is the i^{th} bit of the underlying message which has been encrypted in c, right.

Apart from that, we make no assumption whatsoever about the internal details or internal working of the algorithm A. So, what it models here is that if at all there is a way to identify the i^{th} bit of the underlying plaintext where the underlying message is a random message and encrypted in c, then without even knowing the internal details of the algorithm A, the adversary A' can actually win an instance of the COA experiment. The important thing here is the way reduction has been formulated here is what A' has done here is A' has created a view and by view I mean, the information which has been provided to the algorithm A, namely whatever information the adversarial algorithm A sees, and what adversarial algorithm A sees, it basically sees an encryption of a random *l* bit string and the property of our algorithm A is that it expects an encryption of a uniformly random *l* bit string, and if you give an encryption of a uniformly random *l* bit string with probability half plus non-negligible, it can tell you what exactly is the i^{th} bit of the underlying message.

That is what is the property of the algorithm A, right. So, the way this reduction has been formulated, what the reduction has done is in the reduction A' has created a view which is identical for the algorithm A to launch its attack, right? Because in the COA experiment with probability half, it could be either the message m_0 which could be encrypted in c and the probability half it could be the message m_1 which is encrypted in c, right, and says both these m_0 and m_1 are randomly chosen except they are i^{th} bit. What A' has created for A is actually a view where A is actually seeing an encryption of a random message where the i^{th} bit could be either 0 or i^{th} bit could be 1 with probability 1/2, 1/2 each, right. So, it is very important that A' creates a view which is identical to the view which A expects to launch it attacks, because if the view which A' creates for the adversary A is not identical, then we cannot say that with whatever advantage A can break or identify the i^{th} bit from the ciphertext c, which is almost the same probability A' can identify what message has been encrypted in c.

That relationship we cannot claim if the view which A' has created for A is not identical, which A expects, right. So, let me go back to the previous slide, and in the previous slide, I claimed that the probability with which the adversary A' can win the COA experiment is identical to the probability that the adversary A identifies the i^{th} bit of the random message, which has been encrypted in c. The reason that this equality holds is as I said, when I say that the probability that A could output the i^{th} bit of a randomly chosen message, then with probability 1/2, that randomly chosen message could be the set of all possible *l* bit strings, where the i^{th} bit is 0 and with probability 1/2, it could belong to the set of all possible *l* bit string where ith bit is 1, right? So this probability of adversity A outputting the ith bit is over the random choice of the underlying message. So the probability if I come back to this reduction with probability 1/2, b is going to be 0, and with probability 1/2 be is going to be 1.

That means the c is either going to be an encryption of a random message from the set of *l* bit strings, where the ith bit is 0 or the c could be belonging to the set of random *l* bit strings where the ith bit is 1, and that is what has been forwarded to this algorithm A to identify what exactly is the ith bit. So if indeed, the message m_0 has been encrypted in the challenge ciphertext c, the adversary A is going to output ith bit to be 0 and that ensures that b' = 0, whereas if m_1 has been encrypted in this challenge ciphertext c, that means the ith bit of the message which has been encrypted in this challenge ciphertext is 1, then this adversary A is actually seeing an encryption of a random message whose ith bit is 1, and as per the property of this algorithm with probability half plus non-negligible, it is going to output $m_i = 1$ and with same probability, the algorithm A' is going to output b' = 1. So with whatever advantage, the adversary A' could identify the ith bit of the message which has been encrypted in c with exactly the same probability, our adversary A' could identify what has been encrypted in c.

So that means through this reduction, what we have established is if this adversary A could compute f(m), namely it could identify the ith bit of the underlying message with significant probability better than half, then with almost the same probability, in fact with exactly the same probability, our algorithm A' could actually win the COA game and that will be a contradiction to our assumption to the claim which we are making about our encryption process because we are assuming that our encryption process has indistinguishable encryption.

That means, when I say my encryption process has indistinguishable encryption, that means that no such A' exist which further implies no such A exist, and this is formally established by doing this reduction, right. So, throughout this course, we are going to do proofs like this based on reductions. This is one of the simplest reductions which we have introduced here and you should understand this in a clear fashion, right. So, that brings me to the end of this lecture.

To summarize in this lecture, we have introduced the notion of semantic security in the COA attack model. So, the original definition states that a scheme COA secure if the probability with which adversary could compute some function of the underlying message by seeing a ciphertext is almost the same with which it could compute the same function of the underlying message without actually seeing the ciphertext.

An equivalent version of this definition is the indistinguishability based definition, where the requirement is that if the adversary sees an encryption of a randomly chosen message from a pair of messages where the pair of messages is known to the attacker, then the probability with which it can identify whether it is an encryption of this 0^{th} message or the first message is upper bounded by half plus negligible.

So which can be also stated as the distinguishing advantage of the attacker and distinguishing apart whether the challenge ciphertext it sees in the experiment, in the indistinguishable based experiment belongs to m_0 or m_1, it cannot separate apart except with a negligible probability. We also saw an illustration where we showed actually that if your encryption process satisfies the indistinguishability based definition, then it indeed implies that adversary cannot compute any of the underlying bits of the plaintext. In that illustration, we introduced a reduction-based proof, which are central to cryptography. I hope you enjoyed this lecture. Thank you.