

# **Cryptography and Network Security**

**Prof. D. Mukhopadhyay**

**Department of Computer Science and Engineering**

**Indian Institute of Technology, Kharagpur**

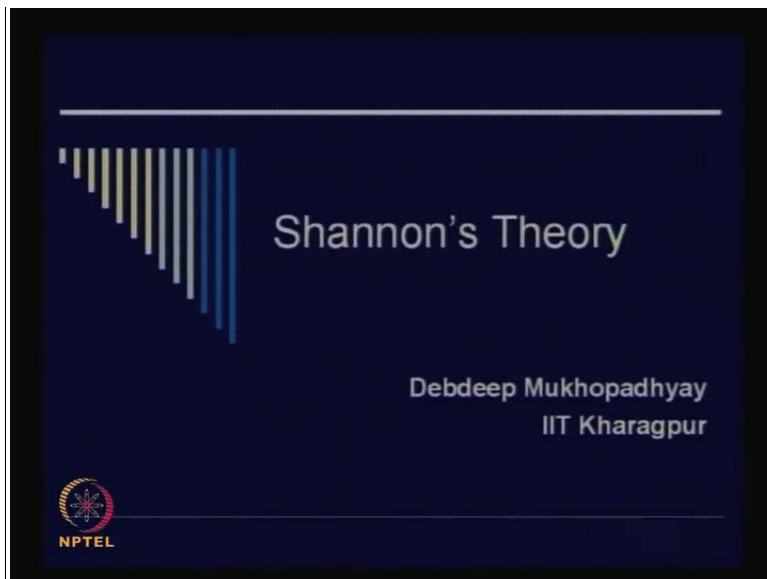
**Model No. # 01**

**Lecture No. # 07**

**Shannon's Theory**

Welcome to this class on Shannon's theory so as I told you in my previous classes that,

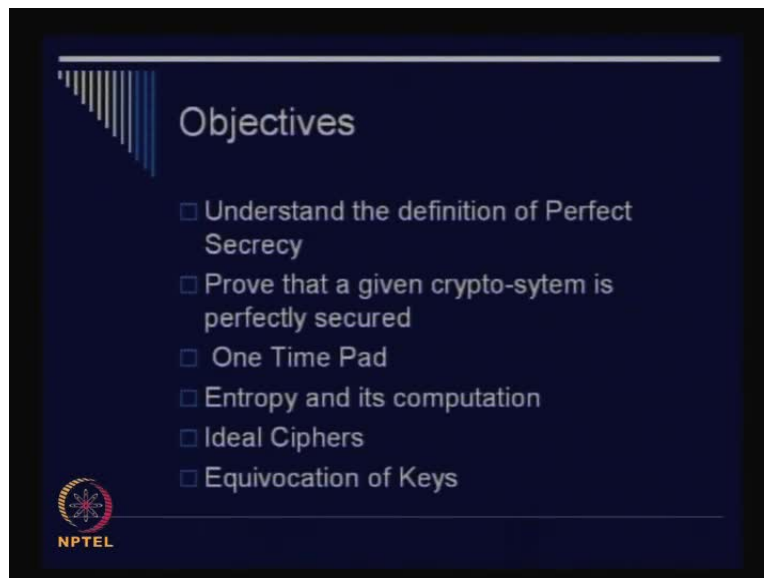
(Refer Slide Time: 00:34)



Shannon's theory is a very fundamental theory in the art or science of cryptology so essentially it was the seminar paper in 1948 and 49 which essentially postulated Shannon theories.

So, they essentially form a cornerstone of what we known as today's ciphers therefore, you will find that whatever we whatever basic formulations or basic properties we find in today's ciphers you can essentially go back and find that those things existed in an old paper in 1948and 1949.

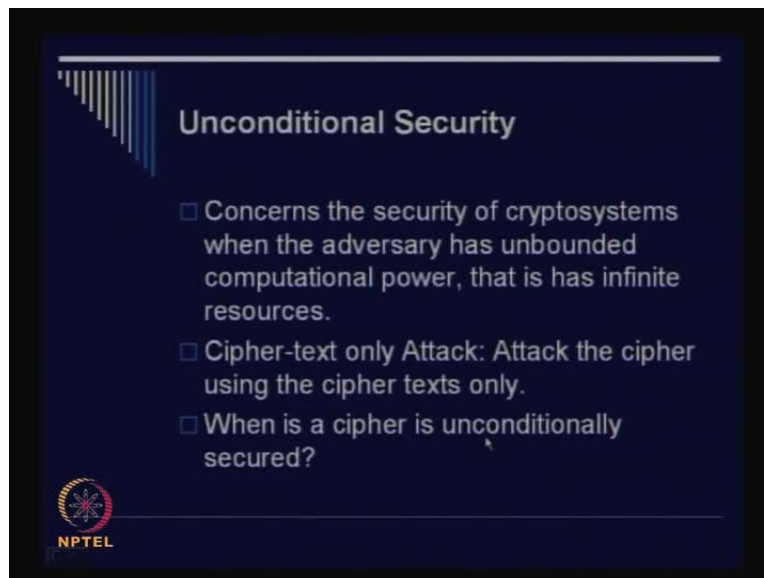
(Refer Slide Time: 01:04)



So we will try to understand some of the concepts in this papers so in today's class the objectives will be as follows so we will try to understand the definition of what is mean by perfect secrecy and prove that at given cryptosystem is perfectly secure so that is our objective that is the given cryptosystem we should be able to essentially find out whether it is perfectly secured are not and then we will see that how to construct or realize a certain kind of perfectly secure ciphers so we call them one time pads or O T P's so after that we will go into a very fundamental instrumental we use for our these kind of proves so it is called entropy.

So we will try to understand what is the entropy and its computations and then we will follow that up with the definition of ideal ciphers and conclude with some a topic called equivocation of keys this i will gradually progress and you will try to understand these concepts.

(Refer Slide Time: 02:09)



By these concepts from a part of Shannon's theory and we will actually conclude with something but, may be not in today's class so first of all what I would like to define what is mean by unconditional security so essentially the idea is that we are essentially considering an adversary who is powerful powerfulness means he has who has got unbounded computational power so the idea is that given a cipher and given a adversary who has got unbounded computational power then whether he or she is able to break a given cryptosystem so that means when we when we talk about unconditional security it concerns the security of cryptosystems when the adversary has unbounded computational power so unbounded computational power means it has got infinite resources so infinite resources means what it has got infinite time it has got infinite space and a question is whether even then whether it can break a given cryptosystem so what do you mean when you say break a cryptosystem.

Obtain by value of the key

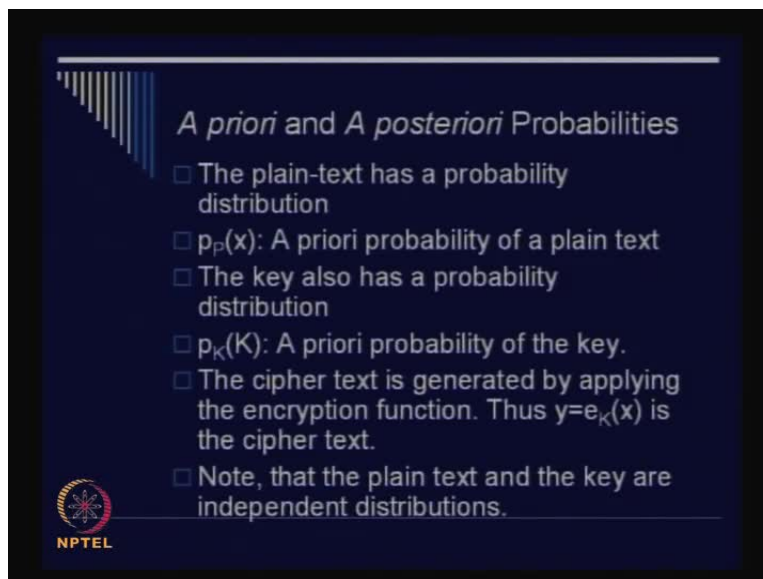
Obtain by value of the key so, essentially we have seen various kinds of attacks we have seen like various classes of attacks so we will be considering that we will considering as hypertext only attack that means the adversary has got access to the only the cipher text and given an unbounded computational power whether the adversary is able to ascertain the value of the key

so, the question always which comes to your mind is when is a given cipher unconditionally secured.

So, in order to understand that we will use some theory of probability and we will try to understand how what is a model of an unconditional adversary and it is an unbounded adversary and try to understand what is the definition of an unconditional I mean unconditional security we also called that perfect security.

that is the basically we have as we have said that the cos curved principle essentially said that the algorithm is open so the security lies only in the value of the key so the value of the key is not known to adversary so the adversary knows the algorithm everybody knows the algorithm the algorithm is in the public domain so what you do not know is the value of the key and a question is where you can ascertain value of the key so therefore,

(Refer Slide Time: 04:27)



*A priori and A posteriori Probabilities*

- The plain-text has a probability distribution
- $p_P(x)$ : A priori probability of a plain text
- The key also has a probability distribution
- $p_K(K)$ : A priori probability of the key.
- The cipher text is generated by applying the encryption function. Thus  $y=e_K(x)$  is the cipher text.
- Note, that the plain text and the key are independent distributions.

NPTEL

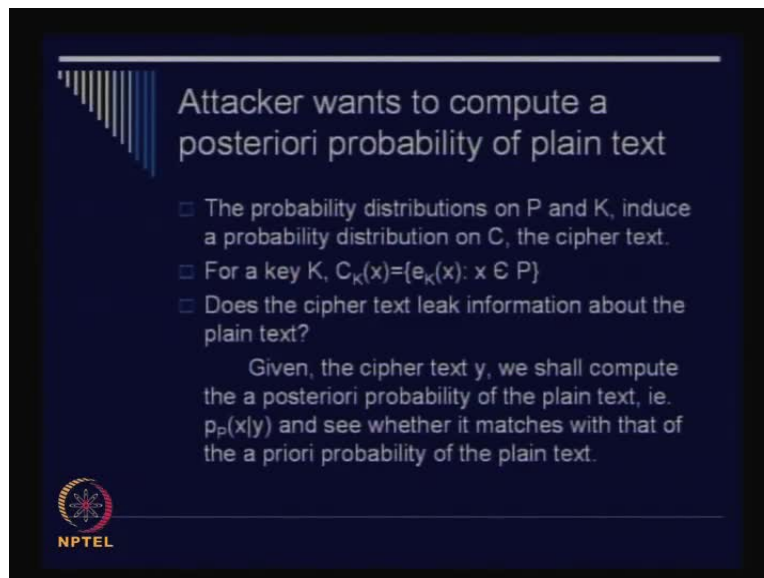
we will use certain times in probability i call that as a priori probability and a posteriori probability so the idea is as follows so whatever encryption you considered is basically a function it is a function from the domain of a plaintext to cipher text therefore, as we have seen in our old classical ciphers they essentially comprised of what they considered for example, let us consider biggest the alphabetic alphabets making of the plain text.

So therefore, how many possible letters or symbols are there there're 26 letters or symbols so that so these 26 letters or symbols have been somehow transformed and converted into a cipher text therefore, all of these letters like a b c d's so and so, on to z has got some probabilities as we have said like in English language e is the most according most figurely according letter.

So, therefore, they have got some probability distribution initially before the cipher text starts and that is what we call as the a priori probability and that is getting essentially therefore, we denote by the term  $p(x)$  so as if you can see here we denote by the term  $p(x)$  a priori probability of a plain text and therefore, this is the probability that is assigned to every letter or symbol of the plain text before the encryption has started so, essentially the key also has got a probability distribution and therefore, now on the in when the encryption function takes place.

Therefore, what happens is that the plain text gets converted into the cipher text so if we consider a cryptosystem where the plain text and the cipher text both comprises of say English letters then the initial frequency of probability distribution of the plain text is getting transformed into a different probability so therefore, that so there is a probability transformation therefore, let us consider like for example, by as we have said that  $p(x)$  denotes a priori probability of the plain text similarly, we have got  $p(k)$  which denotes the a priori probability of the key so the cipher text is generated by applying the encryption function therefore, if we consider that plain text the plain text by  $x$  or denote the plain text by  $x$  and there is an encryption functions  $e$  and based upon the value of the key we know that why  $x$  is getting transformed into  $y$ . So, what we no doubt here is that the both the probability distribution that the plain that is the plain text and the key are independent because when you are choosing the key in general we do not think about what is the value of the plain text therefore, the two distributions that is the plain text distribution and the key distribution they are independent this observation is actually very important so please commit your memory that the plain text distribution and the key distributions are independent we will use that in our future calculations then we obtain the cipher text by applying the value of the plain text and also the chosen value of the key and we obtain as a cipher text in the probability distribution now.

(Refer Slide Time: 07:47)



Attacker wants to compute a posteriori probability of plain text

- The probability distributions on  $P$  and  $K$ , induce a probability distribution on  $C$ , the cipher text.
- For a key  $K$ ,  $C_K(x) = \{e_K(x) : x \in P\}$
- Does the cipher text leak information about the plain text?

Given, the cipher text  $y$ , we shall compute the a posteriori probability of the plain text, i.e.  $p_P(x|y)$  and see whether it matches with that of the a priori probability of the plain text.

NPTEL

What the attacker wants to do is that it wants to compute the a posteriori probability of the plain text so the idea is as follows before going to the complications the idea is as follows

So the idea is like this that is when you obtain the cipher text what you do is that we have got certain values of the symbols so what you do is that you guess the value of the key and then you try to decrypt the cipher text I mean so the moment you decrypt the cipher text what you obtain is the plain text.

So this plain text what you obtain this also has got a distribution and that is called the a posteriori probability now if you are a posteriori probability matches with your applied probability then probably our guess of the key was correct by what if you obtain for example, if you if your guess is wrong then may be when you go back then the plain text that you have obtained does not make any sense so let us I mean if you are not really clear we can make it clear by the example of may be a shift cipher so consider a shift cipher so in shift cipher what did we do every letter was transform by some other letter.

So for example, it could be a shift may be a fixed shift or maybe a not a fixed shift but, we had basically done some transformations therefore, for example, there is a legible or other meaningful English language text consider a paragraph of meaningful English language so when

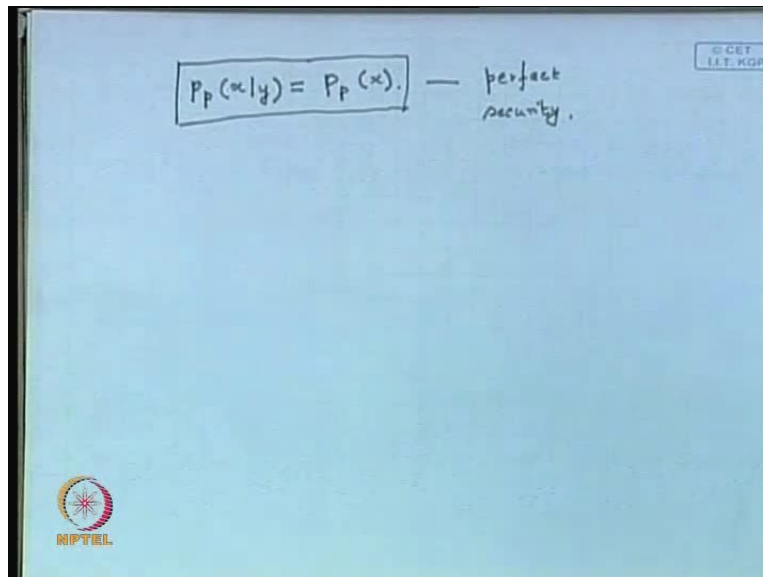
you apply shift cipher gets transformed into another set of alphabets which essentially does not make any sense so what did we do for our crypt analysis we essentially guessed the value of the key and then decrypted back so when we decrypt back we can what we can do is that what we can check what is the original plain text that is our objective also so when you go back we guess the value of the key we go back and obtain the plain text now these plain text or these paragraph that we obtain which is probably the plain text has got a probability distribution but, that is actually based upon the value of the key after the ciphering after the value of the cipher text that is given the value of the cipher text what is the probability distribution of the corresponding plain text.

So these probability distribution is defined as the a posteriori probability and if your guess of the key was correct for a shift cipher then that paragraph could have definitely made a sense you could have been a meaningful text nobody encrypts a meaningless text that is the basic assumption so if some encrypts the meaningless text then I am really not bothered so therefore, the idea is that if that make sense that means essentially in the probability distribution that we obtain in the corresponding text should essentially match with our normal English language therefore, in that case my a posteriori probability matches with the a priori probability and if that is so then the value of the key is correct should you understand the idea behind this therefore, there is a definite idea where Shannon postulated this therefore, essentially it tries to explain mathematically what we did for our crypt analysis.

for example, English language plain text all of us know so the adversary also has a fair amount of idea of that so the idea has an access to that I mean the adversary has an access to that as well it has also access to the encryption function so I am repeating this because it should go into your mind so for example, (Refer Slide Time: 07:47) now consider that in the probability distribution on  $p$  and key  $k$  in this a probability distribution therefore, so this we have understood that the probability distributions on  $p$  and  $k$  that is the plain text and the  $k$  in this a probability distribution on  $c$  which is the cipher text and what is the cipher text we can denote as follows like for a given  $k$   $c = k \oplus x$  is equal to  $e = k \oplus x$  where  $x$  is belongs to the corresponding plain text so this is symbol.

So the question is that so this is the fundamental question does the cipher text leak information about the plain text therefore, given the cipher text  $y$  now what we shall do is the we shall compute the a posteriori probability of the plain text we denote that by  $P_P(x|y)$  so you have understand this so what is this symbol called this is conditional probability therefore, we applied the principles of conditional probability and we can actually compute this value so  $P_P(x|y)$  and see whether it matches with that of the a priori probability of the plain text so if this matches with the a priori probability of the plain text then probably my key is correct therefore, if we would have wanted a proper amount of security are where we would say as a perfect security then we can actually denote this as follows.

(Refer Slide Time: 12:50)

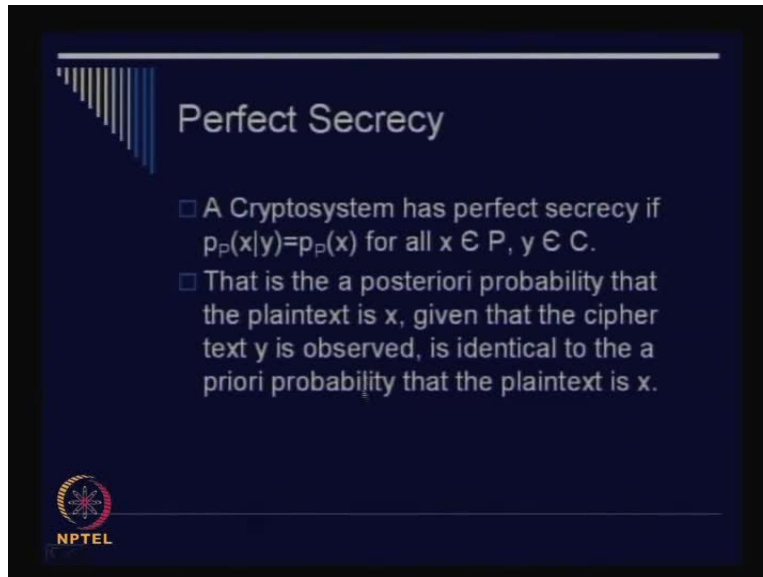

$$P_P(x|y) = P_P(x) \quad \text{— perfect security.}$$

You would have written like  $P_P(x|y)$  that is equal to  $P_P(x)$  so what does it mean so this is the basic idea of a perfect secrecy or perfect security what does it mean.

therefore, I really do not know what to do (Refer Slide Time: 07:47) so the idea is that do not know why this is not shown there so let us continue I am try I mean what I am trying to say it to you is that  $P_P(x|y)$  is equal to  $P_P(x)$  therefore, probably



(Refer Slide Time: 13:53)



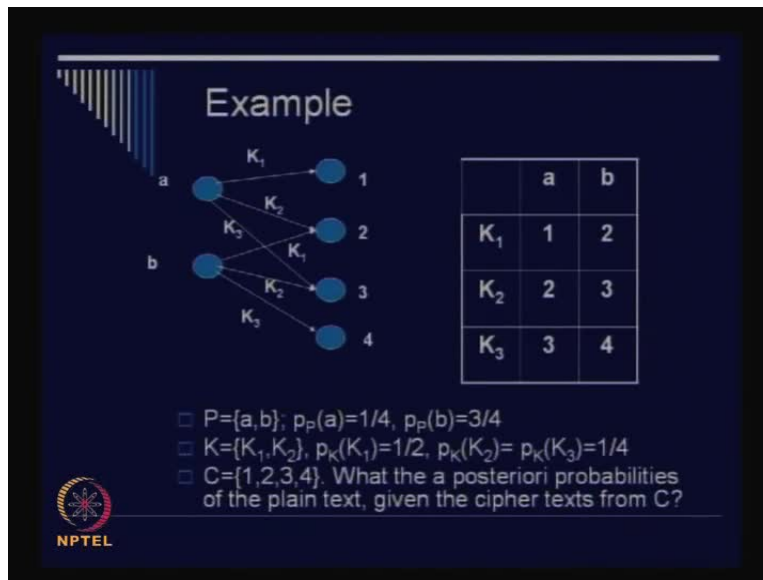
The slide is titled "Perfect Secrecy" and contains two bullet points. The first bullet point states: "A Cryptosystem has perfect secrecy if  $p_P(x|y) = p_P(x)$  for all  $x \in P, y \in C$ ." The second bullet point states: "That is the a posteriori probability that the plaintext is  $x$ , given that the cipher text  $y$  is observed, is identical to the a priori probability that the plaintext is  $x$ ." In the bottom left corner, there is a circular logo with a gear and a star, and the text "NPTEL" below it.

I had written out here somewhere so  $p_P(x|y)$  is equal to  $p_P(x)$  so which means that for all  $x$  given  $y$  belongs to  $P$  and for all  $y$  which belongs to  $C$  this result holds that is the crypto system has got perfect secrecy if this is our case and what does it mean it means that the a posteriori it means that essentially given the value of  $y$  that probability is distribution of  $x$  is indistinguishable from the probability distribution of  $x$  when you are not given in any value of  $y$  so which means what which means that a cipher text  $y$  is not giving you any additional information so you see that slowly you are trying to understand the meaning of information or the meaning of uncertainty so information and uncertainty are quite related so we are trying to say how those things are formally destroyed mathematically destroyed.

I you are not understood so there was he so what I am trying to say is that suppose consider an symbol shift cipher and consider that this is the probability distribution of  $x$  so that is I have describe by  $p_P(x)$  what is the normal English language description now you consider that you have done a shift cipher so which means what you have got a  $y$  cipher text now I am saying to you that given this distribution of  $y$  whether the now what you can do is that you can the guess value of the key and obtain like the plain text so that is nothing but, so the probability distribution now that you have obtain is  $p_P(x|y)$  so if these two things so that is what I have trying to say if these two things match then essentially this means that this  $y$  is not giving you

any extra information because even then it is still the same so it is not giving you any additional information so,

(Refer Slide Time: 15:47)



Let us talk lets I think get I mean I think this will gets clear (Refer Slide Time: 07:47) with the help of an example therefore, let see one example therefore, consider this is the cryptosystem obvious just with the sort of diagrammatic representation of the cryptosystem.

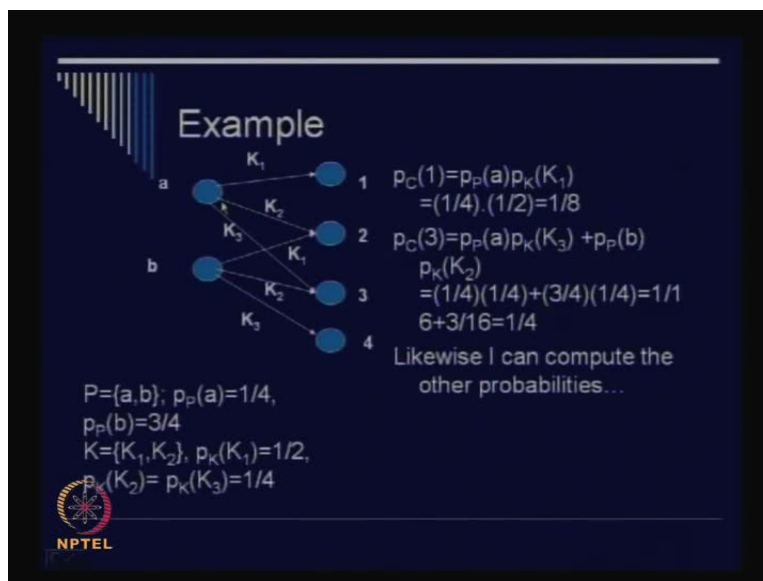
So, you see that this is a very simple transformation which says that your plain text comprises of the letters a and b and your cipher text comprises of the letters or symbols 1 2 3 and 4 so what is the transformation the transformation says that you can start with a and if your k is k 1 then you go to 1 if your k is k 2 then you go to2 if k is k 3 you go to 3 similarly, the other transformations so what you see is that if your plain text is b then also you get some other mappings like 2 3 and 4 therefore, this is the corresponding mapping that we are concerned with now let us try to understand sudden things about what we are just described so, for example, consider that your plain text which forms the set of p equal to a b has got a probability distribution.

So, which means you're a has a probability distribution a b has a probability distribution so essentially follows the probability distributions therefore, a has a probability of occurring and b also has a probability of occurring so what is the probability of occurring of a it is say 1 by 4 so

this is given so what is the probability of occurrence of b it is 3 by 4 note that I mean the probability of occurrence of a plus the probability of occurrence of b is equal to unity so this is obvious because a and b are the only possible texts symbols so what is the value of k the k could be k 1 k 2 and probably k 3 also, this is the mistake here therefore, so consider that suppose there's key called k 1 and the probability distribution of k 1 is half and the probability distribution of k I mean the probability of occurrence of k 2 and k 3 as same so what is that it is equal to 1 by 4 so what is the cipher text the cipher text is 1 2 3 and 4.

Now the question is what is the a posteriori probabilities of the plain text given in the cipher text from c so which means that we have been provided the corresponding cipher text and the question is what is the a posteriori probability of the plain text so which means what is the probability of occurrence of say a or b once you have been given say the cipher text is 1 or 2 or may be 3 or 4 so you will consider one such case now this encryption function is often represented by this table also, you see that a is the plain text k 1 is the key then your cipher text is one similarly, you can also check other such mappings is this clear.

(Refer Slide Time: 18:47)



So, I will keep this figure and I will try to compute the value of the a posteriori probability as follows therefore, first of all let us try to understand what is the probability that one occurs at the cipher text.

So, what is that probability so you see that one could have occurred only from a in this matrix so that means we would like to multiply because you already told you whether plain text distribution and the key distribution are independent distributions so in order to obtain the probability of 1 we can simply multiply the probability that a occurrence of a multiply with the probability of the key k 1 so what is the probability of a it is equal to 1 by 4 and the probability of occurrence of k 1 is half so multiply 1 by 4 with half and obtain 1 by 8 so that is the probability of occurrence of 1 similarly, we can also obtain the probability of occurrence of three it is slightly more complicated because three can occurred from a as well as b but, in this case you see that this a and these b are two exclusive cases therefore, we can apply the theory of or theorem of a priori probability and what we can do is that we can multiply the probability of occurrence of a with the probability of occurrence of k 3 and add that with the probability of occurrence of b multiply by the probability of occurrence of k 2.

So, this is symbol so what we can do is that we can just simply multiply 1 by 4 with 1 by 4 add that product with the 3 by 4 and 1 by 4 and this works to 1 by 4 so this is actually 1 by 16 plus 3 by 16 not 1 by 1 so it is equal to 1 by 4 so likewise I can also compute the other probabilities so you see that we can obtain the probability of the cipher texts as follows but, our question is what our question is to find out the a posteriori probability which means is given the cipher text is 2 I would like to compute what is the probability of occurrence of a so this is quite simple likewise if i say you for example, that the cipher text is 1 and ask you what is the probability of occurrence of b what is the answer it is 0 you can easily compute so you see that does it match with the probability of occurrence of b no which means that it is not a.

Perfect cipher

(Refer Slide Time: 21:16)

**Example**

$P = \{a, b\}; p_p(a) = 1/4, p_p(b) = 3/4$   
 $K = \{K_1, K_2\}, p_K(K_1) = 1/2, p_K(K_2) = p_K(K_3) = 1/4$

- $p_p(a|1) = 1; p_p(b|1) = 0$
- $p_p(a|2) = ?$
- The '2' can come when the plain text was 'a' and the key was 'K<sub>2</sub>' or when the plain text was 'b' and the key was 'K<sub>1</sub>'
- Given '2', we need to compute the probability that it came from 'a'.
- Is it that of choosing K<sub>2</sub>? No.

NPTEL

Perfect cipher so this is the symbol check but, we will try to compute little bit complex situation which says that what is the a posteriori probability of a given that cipher text is 2 we will just try to understand this complicated thing although we are easily understood that the this is not a perfect cipher why because we have to essentially go to the generalization so generalization are always complex equation so in order to understand that it is interesting to work with a simple example so let us consider this for example, that suppose your  $p_p(a)$  is given 1 is 1 and your  $p_p(b)$  given 1 is 0 so, this we of all already understood I guess that is your if your plain text is 1 your cipher text can be a and that occurrence probability is actually 1 and if your plain text is 1 then the occurrence of b as the plain text is actually 0 so you would like to compute this also and what is the a posterior probability that a accords as a plain text given 2 as a cipher text so that this is two can actually come from two plain text it come from a as well as it can come from b so the two can come when the plain text was a and the key was k 2 or when the plain text was b and the key was K 1.

K 1 so another question is given to we need to compute the probability that it came from A is it that of choosing k 2 is that probability is same as that of choosing k 2.

No because we see that there are other mappings where k 2 has been chosen where 2 is not the result for example, this one in series mapping the k 2 is chosen but, actually have ended up with three as a cipher text so it is not equal to that of choosing of the k 2 so,

(Refer Slide Time: 23:03)

**Example**

Given '2', we need to compute the probability that it came from 'a'.

The '2' can appear with a probability:

- by having 'a' as the PT and  $K_2$  as the key:  $(1/4)(1/4)=1/16$
- by having 'b' as the PT and  $K_1$  as the key:  $(3/4)(1/2)=6/16$

$p_P(a|2)=(1/16)/(7/16)=1/7$

$P=\{a,b\}; p_P(a)=1/4, p_P(b)=3/4$   
 $K=\{K_1,K_2\}, p_K(K_1)=1/2, p_K(K_2)=p_K(K_3)=1/4$

NPTEL

Therefore, how do we compute that therefore, let us see I mean for example, these two can appear with a probability as which we can work out as follows by having a as a plain text and k 2as a key therefore, A is the corresponding plain text and k 2 is the value of the key so what is the probability so we know that we have to we can multiply the probability of the plain text and the key because they are independent so it is product of 1 by 4 and 1 by 4 and it works to 1 by 16 you understand that because product of occurrence of a is 1 by 4 and k 2 as the key is also 1 by 4 so multiply them it is 1 by 16 the other case could be by having b as the plain text and k 1 as the key you see that the other occurrence the other chance could be that b was the plain text and k 1 was the key thus k 1 is chosen so what is the probability of occurrence of b it is 3 by 4 and what is the probability of occurrence of k 1 it is half so you multiply them and you obtain 6 by 16

So what is the total probability that two can occur it is 7 by 16 so we have basically broken up the possibilities into two cases so what is our desirable case now our desirable case is that two has occurred as a cipher text and it has occurred from a so which is the first event you see that by having a as the plain text and k two as the key what we have done is that what we have obtained two from the plain text a and that is precisely what we want what is the probability distribution of a given that two is your cipher text so what you do is that we divide 1 by 16 by 7 by 16 and we obtain the value 1 by 7 so, I guess you understand began now generalize this also like consider any set of plain text and any corresponding set of cipher text and any set of keys for of course, so we would like to generalize this idea of a posteriori probability and this we can do by the help of this equation.

(Refer Slide Time: 25:06)

Generalization of the Example

$$P_P(x|y) = \frac{P_P(x) \sum_{K: x=d_K(y)} P_K(K)}{\sum_{K: y=C(K)} P_K(K) P_P(d_K(y))}$$

NPTEL

(Refer Slide Time: 25:42)

**Example**

$P=\{a,b\}; p_P(a)=1/4, p_P(b)=3/4$   
 $K=\{K_1,K_2\}, p_K(K_1)=1/2, p_K(K_2)=p_K(K_3)=1/4$

- Given '2', we need to compute the probability that it came from 'a'.
- The '2' can appear with a probability:
  - by having 'a' as the PT and  $K_2$  as the key:  $(1/4)(1/4)=1/16$
  - by having 'b' as the PT and  $K_1$  as the key:  $(3/4)(1/2)=6/16$
- $p_P(a|2)=(1/16)/(7/16)=1/7$

NPTEL

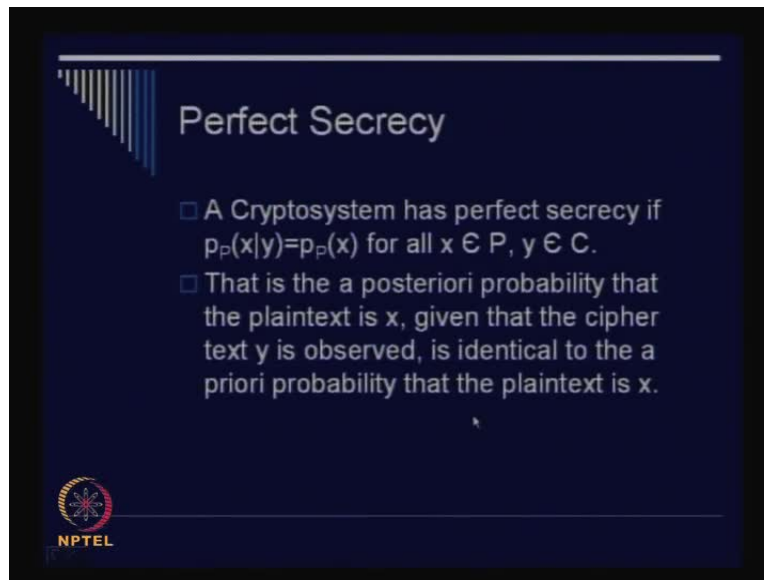
Do you understand this equation so you see that what you says as follows so it is basically tries to compute the value of  $p_P(x|y)$  given  $y$  so this is what exactly we are doing in our example so the denominator is all possible ways how you can actually obtain  $y$  because  $y$  is the cipher text so you remember in the previous case what did we do we essentially took that corresponding cipher text and decrypted them by all possible values of the key and went back to the corresponding plain text and then multiplied with the probability of that plain text you see what I am saying what I am saying is that in order to compute this total probability that is of occurrence of two what we did is that we took for example, any key like for example, choose  $k_1$  and decrypted this two with  $k_1$  went back to the corresponding plain text and multiplied with the probability of the plain text and again choose  $k_2$  multiply I mean obvious of quality of  $k_2$  and then decrypt it and go back to  $a$  and multiply with the probability of  $a$ .

Similarly we can do it for all possible keys so that is exactly done here so in this case you see that you take you take the corresponding cipher text  $y$  you decrypt them by all possible keys multiply with the probability of the I mean the corresponding the plain text and multiply with the probability of the key and then the sigma of all possible of such keys and in interesting case is that is the fact that your that your  $x$  has occurred as the plain text and you essentially multiply them with all possible keys which essentially takes you from the value of  $x$  to the value of  $y$  so



essentially it is therefore, this is exactly what we have seen in the previous example is this generalization clear to us.

(Refer Slide Time: 27:03)

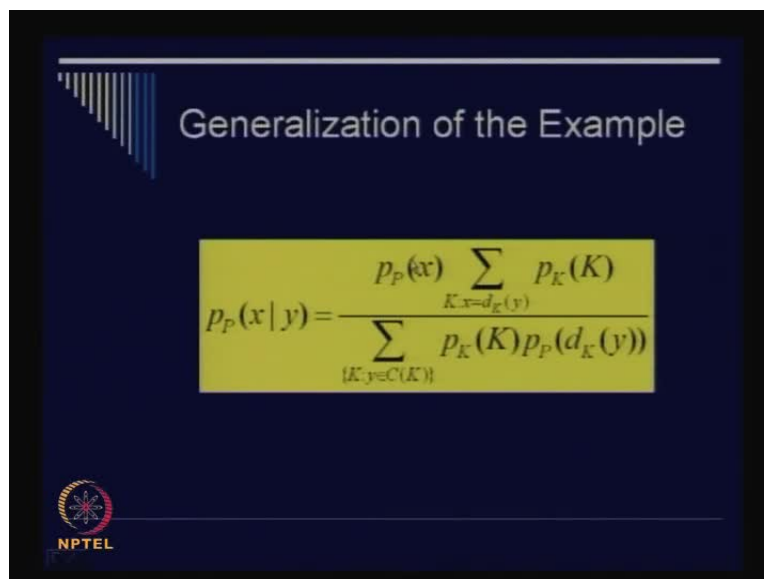


**Perfect Secrecy**

- A Cryptosystem has perfect secrecy if  $p_P(x|y) = p_P(x)$  for all  $x \in P, y \in C$ .
- That is the a posteriori probability that the plaintext is  $x$ , given that the cipher text  $y$  is observed, is identical to the a priori probability that the plaintext is  $x$ .

NPTEL

(Refer Slide Time: 27:31)



**Generalization of the Example**


$$p_P(x|y) = \frac{p_P(x) \sum_{K: x=d_K(y)} p_K(K)}{\sum_{\{K: y=C(K)\}} p_K(K) p_P(d_K(y))}$$

NPTEL

Yes therefore, so now essentially we know how to compute the value  $p_P(x|y)$  given  $y$  and essentially I think I have already defined but, I mean given you the idea this is the formal definition it says that a crypto system has got perfect secrecy if the value of  $p_P(x|y)$  is equal

to the value of  $p(p|x)$  for all  $x$  given belongs to  $p$  and all for all  $y$  belongs to  $c$  and now the value of  $p(p|x)$  we already know we can engage this previous equation to the calculate the value of  $p(p|x)$  given  $y$  because you see that on the hand side everything is known to us we know the value of the key we know the value of the  $I$  mean the probability distribution of the plain text as well as the  $I$  mean the probability distribution of the key and therefore, we can use this equation because we have also know the mapping that we know the decryption function so we can use this formula to calculate the value of  $p(p|x)$  given  $y$  and check whether this value of  $p(p|x)$  given  $y$  matches with the probability of  $p(p|x)$  if  $p(p|x)$  given  $y$  matches with  $p(p|x)$  then we have got perfect security.

(Refer Slide Time: 28:28)



**Shift Cipher has perfect secrecy**

- Suppose the 26 keys in the Shift Cipher are used with equal probability  $1/26$ . Then for any plain text distribution, the Shift Cipher has perfect secrecy.
- Note that  $P=K=C=Z_{26}$  and for  $0 \leq K \leq 25$
- Encryption function:  $y=e_K(x)=(x+k) \bmod 26$

So that is the idea so that informally it means that is a posteriori probability that plaintext is  $x$  given that  $y$  the cipher text  $y$  is observed is identical to the a priori probability that the plaintext is  $x$  so that is the definition of perfect security so then let see when a some examples where we can have perfect secrecy so consider a simple example of a shift cipher so in a shift cipher what we did in our  $i$  mean what we have seen in our little bit not so simple shift cipher if you remember we had taken a and we have mapped to any of the possible outputs so there are 26 possibilities for the first letter the second letter  $b$  maps to any of the symbols except what a have been transformed to so there were 25 such possibilities so if we multiply keep on multiply in such a fashion then the size of the key is equal to 26 factorial we have seen that.

So what we will try to see here is that if each possible mapping is chosen randomly then shift cipher achieves perfect secrecy so we will try to use our previous formulation and you will try to establish this fact if suppose the 26 keys in the shift cipher are used with equal probability so what means that that means that each plain text I mean whether each key occurrence of each key has got a probability of one by 26 therefore, any plain text distribution the shift cipher has got perfect secrecy so which means that actually we are not bothered also what is the plain text distribution and we can still achieve perfect secrecy so do you see how so you see that we have  $p_K$  and  $c$  can be set to form the set  $Z_{26}$  so  $Z_{26}$  we have seen that if the set of integers from 0 to 25.

So, the  $k$  therefore, immediately understand lies between 0 and 25 and your function is very simple it is just  $x$  plus  $k$  mod 26 so idea is that if for each symbol of the plain text if we choose the value of the key at random that is any mapping you can just choose arbitrarily then essentially if and if you obtain the corresponding value of the cipher text so then the question is whether this gives you perfect secrecy so which means that when you choose so it means it means we should try to understand this fact that what we are saying is not such scenario in which for all the plain text we are using the same key it means that for each symbol that each symbol which you want to transmit secretly you have to essentially choose the value of the key also at random.

(Refer Slide Time: 32:21)

**Perfect Secrecy**

$$P_P(x|y) = \frac{P_P(x)P_C(y|x)}{P_C(y)}$$


$$P_C(y) = \sum_{K \in Z_{26}} P_K(K)P_P(d_K(y))$$

$$= \sum_{K \in Z_{26}} \frac{1}{26} P_P(y-K) = \frac{1}{26}$$

$$P_C(y|x) = P_K(y-x \text{ mod } 26)$$

$$= \frac{1}{26}$$

*Hence Proved*

 NPTEL

So for example, there is a letter like say ababraka dabra and if you have been encrypts a b c like so on then for a you have been choose the corresponding key at random for b also you have to choose the key at random for c also you have to choose the key at random so it does not mean that for the entire plain text we are using the same key so for each symbol we are choosing a value of key at random so what is the what do you see is the practical implication practical implication is you see the practical problem the practical problem is that for each symbol you are choosing a key at random so in order to decrypt also you require the same key so you immediately understand that you have to transmit lots of keys and an it is actually in this case the same as that number of the plain text that you want to transmit so if how to transmit the key secretly then why not do it do the same for the plain text also so the even though you obtain a perfect secrecy and also in a perfect secrecy but, that is not practical so, we will try to first establish why it is a perfectly secured cipher and later on go into a another case where we actually obtain a perfect secrecy it is called a onetime pad and then talk about its practical problems so this is the mathematical formulation it says that  $p(p, x \text{ given } y)$  we have already seen this is equal to  $p(p, x)$  multiplied with  $p(c, y \text{ given } x)$  divided by  $p(c, y)$  these are same thing as what we have seen in the previous equation so it is a so the denominated is the probability that y has occurred as the cipher text and the numerator says that x is occurred as the plain text and that has been multiplied with the probability that y has occurred given x is a plain text so can you understand what is the value of this  $p(c, y \text{ given } x)$  it is the same as the probability that the key has been chosen and that key is equal to  $y \text{ minus } x \text{ mod } 26$ .

So, you see that when we are computed the value of  $p(p, x \text{ given } y)$  that was not equal to the value of the probability of the key we discussed that but, in this case its equal because if you choose the value of the key to be  $y \text{ minus } x$  and if your plain text is x then obviously your cipher text is y therefore, this probability you can denote by this symbol it says that you can actually multiply with the corresponding value of the key so you see that  $p(c, y \text{ given } x)$  is equal to the value that the is equal to the probability that I mean we have to basically choose the fact that the key is equal to the  $y \text{ minus } x \text{ mod } 26$  what is the probability that the key is equal to  $y \text{ minus } x \text{ mod } 26$  and what is the probability that is  $1 \text{ by } 26$  because we have assumed that each key is being chosen at random.

Similarly we obtain the value of the corresponding value of the  $y$  occurring as the cipher text so what is the probability that  $y$  occurred as cipher text so you choose all possible keys and you decrypt your  $y$  by using your decryption function multiply that with the corresponding I mean corresponding value of the key been chosen so what is the value of the key 1 by 26 and what is the value in I mean what is  $d_k d_k y$  I mean what this is some probability so you can keep it  $p_p y$  minus  $k$  but, when you take the  $\sum_{k=1}^{26} p_p c_y$  comes out and you have got a summation of our probabilities so what is that equal to its equal to 1 so you have got 1 by 26 now you see that the result is established because  $p_p c_y$  given  $x$  cancels with  $p_c y$  both of them are equal to 1 by 26 so if these two things canceled then you have got  $p_p x$  given  $y$  is equal to  $p_p x$  so what does it mean.

(Refer Slide Time: 35:44)

**Theorem**

- Suppose  $(P, C, K, E, D)$  be a cryptosystem, where  $|K|=|C|=|P|$ . The cryptosystem offers perfect secrecy if and only if every key is used with probability  $1/|K|$ , and for every  $x \in P$  and every  $y \in C$ , there is a unique key, such that  $y = e_k(x)$ .
  - Perfect Secrecy (equivalent):  $p_C(y|x) = p_C(y)$
  - Thus if Perfect Secret, a scheme has to follow the above equation.

NPTEL

You have got perfect secrecy why and because here a priori probability of  $x$  given the value of  $y$  matches with you're a priori probability of  $x$  so if you can do this for shift ciphers we can do it for any given cipher we can basically what you have to do is that we have to compute these values of probabilities and try to see whether we get a match so this was just as very simple example so you can if you see your Stinson's book there are several exercises given and I will also give you one as an exercise so you can just practice and if you can do it for one probably you can do it for others as well so basically now let us try to think of a theorem or let us try to establish a theorem it says that any encryption function you can or any cryptosystem essentially

can be denoted by a 5 tuple you have got  $p$   $c$   $k$  so what is that plain text cipher text and the key and your encryption function and your decryption function if these five things are defined then you have defined a cryptosystem.

So let us consider the case that the size of the  $k$  size of the  $c$  and the size of the plain text are all the same that is your cardinality of your key and your cardinality of the cipher text and cardinality of the plain text are the same so the cryptosystem offers perfect secrecy if and only if every key is used with probability of  $1$  by modulo  $k$  so in your previous example there were how many keys there were  $26$  keys and the occurrence of each key was  $1$  by  $26$  and that lead to the fact that we had a perfect secret system perfectly secret system so and for every  $x$  which belongs to  $p$  and for every  $y$  belongs to  $c$  there is a unique key so that means that given  $x$  and  $y$  there is a unique key which takes  $x$  and maps it  $y$ .

So we have to establish this fact that is for every  $x$  and  $y$  you choose or every tuple that you form with  $x$  and  $y$  that is at given plain text and a given cipher text you have a unique value of key which defines this so we will try to establish this fact so I will although I will not prove but, you can easily understand that an equivalent definition of perfect secrecy will be  $p(c|y)$  given  $x$  is equal to  $p(c|y)$  I leave it to you as an exercise you can actually prove it by theorems of conditional probability its very simple straight forward therefore, it says that the cipher text probability of  $y$  is a same as the probability of  $y$  given  $x$  as the corresponding plain text so thus if perfectly secret a scheme has to following the above two equations and both of them are equivalent if we can proof any one of them the other one is proved equivalently.

(Refer Slide Time: 39:01)

## Cryptographic Properties

- $p_C(y|x) > 0$
- This means that for every cipher text, there is a key,  $K$ , st.  $y = E_K(x)$
- Thus  $|K| \geq |C|$ . In our case,  $|K| = |C|$
- Thus, there is no cipher text,  $y$ , for which there are two keys which take them to the same plaintext.
- There is exactly one key, such that  $y = E_K(x)$

NPTEL

So, we observe from fact so we see that since  $p_C(y|x)$  given  $x$  is equal to  $p_C(y)$  that is if you see this equation it means that if you fix a value of  $x$  for example, as a plain text then your  $p_C(y|x)$  is equal to  $p_C(y)$  and your  $p_C(y)$  is greater than 0 because if that  $y$  I mean the occurrence of  $y$  the probability of occurrence of  $y$  would have been 0 then  $y$  wouldn't have appeared in the cipher text set so that means for some cipher text at least for some plain text at least we have that corresponding cipher text so immediately which means that this leads to the fact that  $p_C(y|x)$  is also greater than 0 so what I am trying to say is that since  $p_C(y)$  is greater than 0 and if you have a perfectly secured cipher then  $p_C(y|x)$  given  $x$  being equal to  $p_C(y)$  this is also greater than 0 therefore,  $p_C(y|x)$  is greater than 0.

So, which means that for every cipher text there is a key  $k$  such that  $y$  is equal to  $E_k(x)$  so, which means that for every cipher text there is a key such that  $y$  is equal to  $E_k(x)$  so if you think in the terms of your diagram you could have understood that in this in the corresponding cipher text said there is no single point which is not being mapped that is not any isolated point which is not mapped to any plain text therefore, your plain text comprises of what  $x_1, x_2$  and so on symbols your cipher text also comprises of  $y_1, y_2$  and so on symbols so that is not a scenario where there is a cipher text symbol say some say corresponding consider any  $y_i$  for example, which is not mapped to any plain text  $x_i$  do you understand that.

So, that means that for every cipher text there is a key there is one key at least which will take it to that  $x$  be that any  $x$  so what does it mean it means that the size of the key set is obviously at

least great equal to the value size of the cipher text it can be greater than that also but, it is at least equal therefore, what I am trying to say is that for perfect secure ciphers the size of the key or the cardinality of  $k$  is more than the cardinality cipher text so you can understand again I mean again retaining that the fact what I am trying to say is that for every cipher text  $y$  there is a key  $k$  which defines this mapping so that means there has to be the number of the keys has to be at least equal to the numbers of the cipher text for every cipher text symbol you have a separate key which will define this mapping.

So number of keys is obviously equal to the size of the  $y$  size of  $y$  it can be more than that also but, at least equal therefore, the size of the key is greater than the size of the  $c$  in our case consider that a size of  $k$  and a size of  $c$  are the same so thus there is no cipher text therefore, now we can actually understand from here also that there is no cipher text  $y$  for which there are two keys which take them to the same plain text why and because your what is your size what is the how can you define your cipher text set  $c$  you can define your cipher text set  $c$  by the set that you basically choose any plain text  $x$  and you transform that by  $e_k x$  and you obtain the cipher text so basically that suppose if I fix the value of  $x$  I could have chosen all the possible values of keys and  $i$  would have obtain the entire cipher text set.

So that we know that if we say that so this is your size of the cardinality of this set is the same as that of your cipher text now if this is equal to the size of the value of the set of  $k$  that is it is as same as size of  $k$  it means that there are no two keys which are distinct like  $k_1$  and  $k_2$  such that  $e_{k_1} x$  is equal to  $e_{k_2} x$  because in that case one cardinality would have been less one sets one of the sets cardinality would have been lesser than the cardinality of  $k$  and what we have said is that if you assume that the size of  $k$  and the size of  $c$  are same then essentially there is no two keys which are distinct which will take them to the same plain text thus there is exactly one key which defines the mapping from  $x$  and  $y$  so therefore, I mean this is the basic definition so we see certain definitions property of the idea of a perfect secure cipher.



So, what are the those definitions one important point was that you're a posteriori probability of your plain text matches with the a priori probability the second definition was that equivalent definition was that  $p(c|y \text{ given } x)$  was equal to  $p(c|y)$  so that is also same thing that the a priori probability of your cipher text is the same as the a posteriori probability of your cipher text and the third important thing which we see is that if your size of plain text key and cipher text are the same then essentially for every plain text  $x$  and for every cipher text  $y$  for any ordered pair like this that is the unique key will defines this mapping so there are no two such keys like  $k_1$  and  $k_2$  which will take  $x$  for example, as a plain text and we will map it to the same cipher text  $e_{k_1}$  where  $k_1$  is not equal to  $k_2$  we cannot have the scenario where  $e_{k_1}(x)$  is equal to  $e_{k_2}(x)$  because if that be the case then your size of cipher text would have been lesser than that of the size of the key but, what we have said is the size of the cipher text is equal to the size of that is this part clear so we will stop at this point and essentially break for a few minutes and come back with the definition of something which we called as the one time pads.