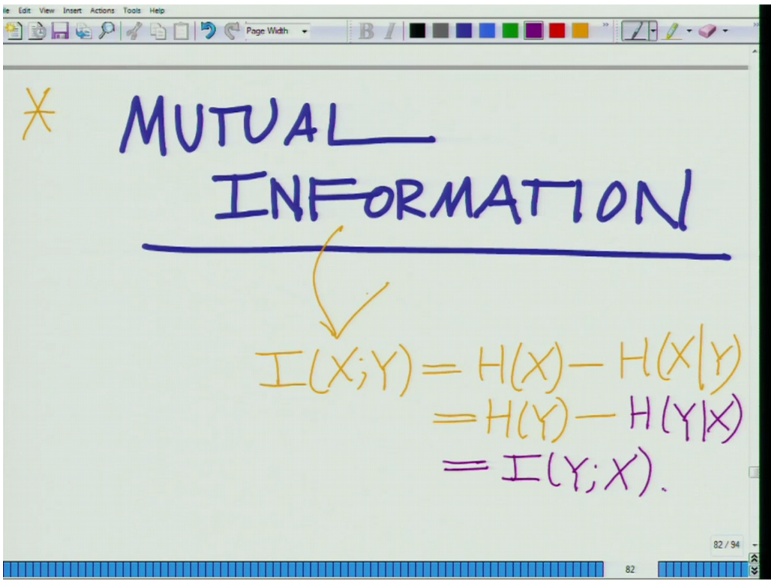
Principles of Communication Systems - Part II
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 34
Simple Example of Mutual Information, Practical Example of Mutual Information - Binary Symmetric Channel

Hello, welcome to another module in this massive open online course. So, we are looking at mutual information, correct.

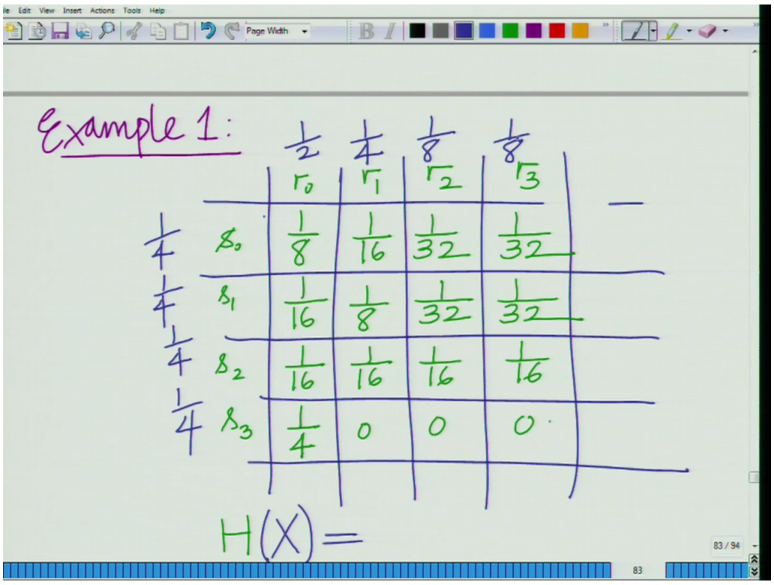
(Refer Slide Time: 00:19)



Which we said as a very key role in the study of communication systems mutual information and if we look at this; this is defined as I X Y equals H X minus H X given Y which is equal to H Y minus H Y given X and by symmetry this is also the mutual information between Y comma X.

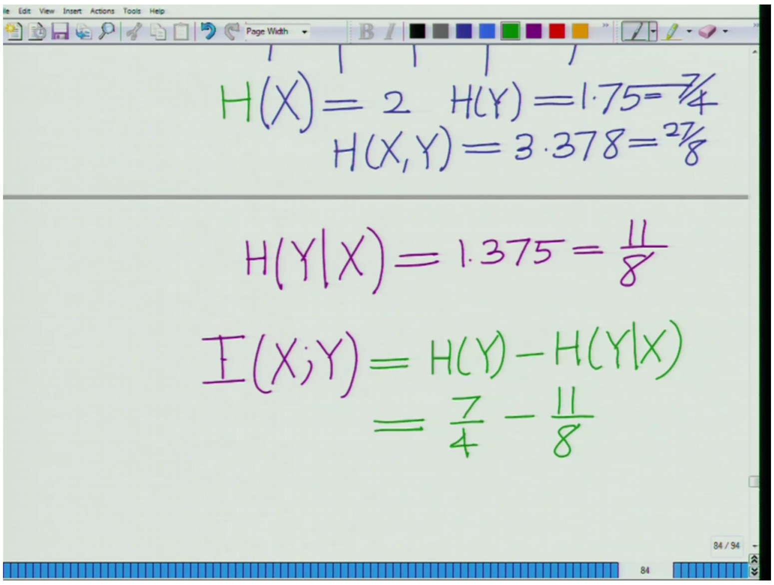
Now, let us go back to our example that we had seen the simple example that we had seen earlier with the joint probabilities I am going to illustrate 2 examples to complete the mutual information. So, first let us go back to our example 1 with the simple table for the joint probabilities.

(Refer Slide Time: 01:29)



So, from example from this example you can see. So, if you look at this s 0 s 1 s 2 s 3 r 0 r 1 r 2 r 3 the various probabilities are the joint probabilities are 1 by 8 1 by 16 1 by 32 1 by 32 1 by 16 1 by 8 1 by 32 1 by 32 1 by 16 1 by 16 1 by 16.

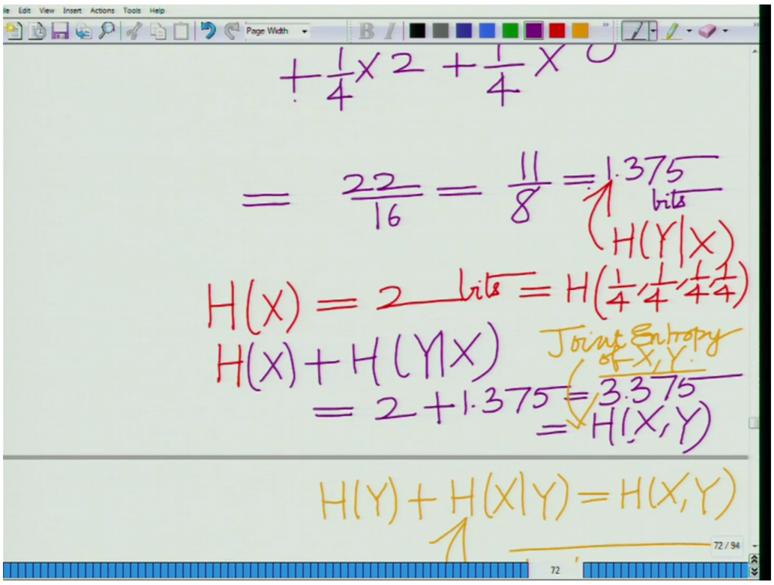
(Refer Slide Time: 02:37)



And if you remember for this example we had seen that well we had seen the following quantities we are calculated the based on the marginal probabilities H X equals. So, the marginal probabilities are H as all the symbols have equal probabilities 1 by 4 1 by 4 1 by 4 1 by 4 X can take either symbols s naught s 1 s 2 s 3 and why as the following probability as half 1 by 4 1 by 8 1 by 8.

So, H X is basically 2 bits H Y equals 1.75 H X comma Y equals 3.375 this is equal to well this is equal to 7 by H Y equals 7 by 4 H X if I am not mistake and that is equal to 27 by 8 of course, we have also calculated H Y and H Y given X. Remember if you remember we add also calculated if you go when we define the conditional entropy we had also calculated H Y given X `and H Y given X you can see here.

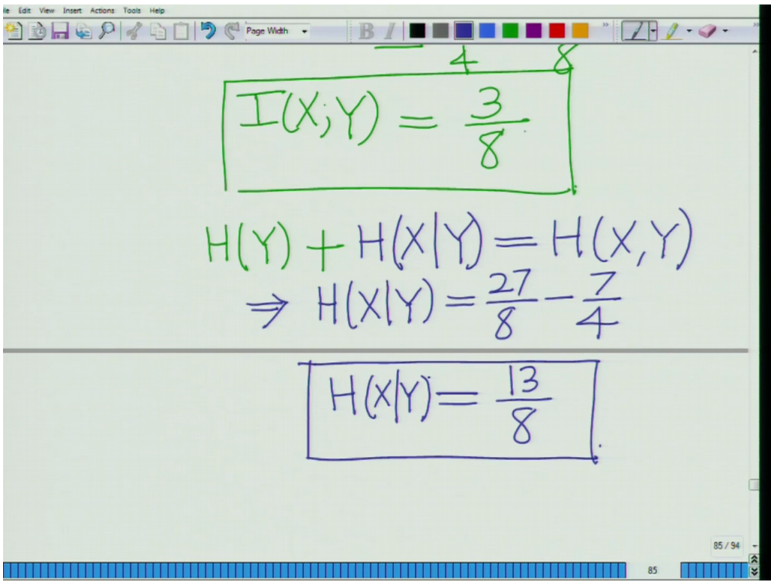
(Refer Slide Time: 04:00)



This is the value of H Y given X that is 1.375 bits 1.3 H Y given X is 1.375 that is 11 by 8, correct.

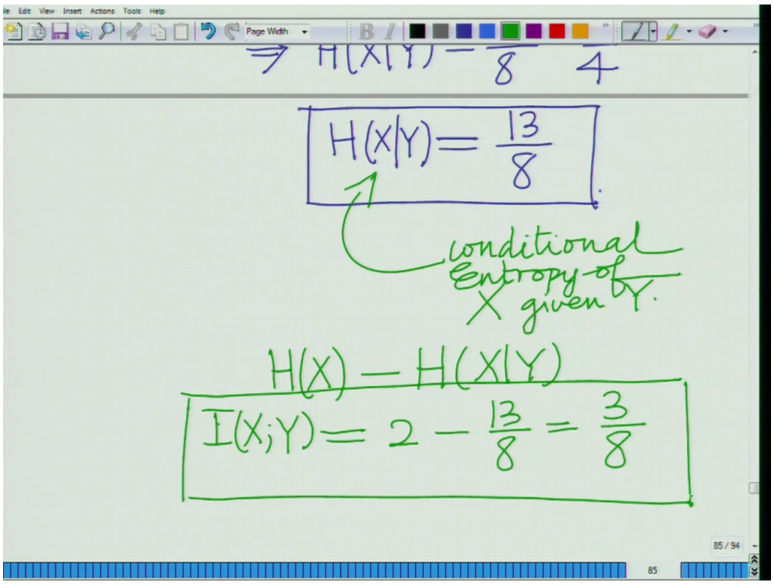
So, H Y given X equals 1.375 that is equal to 11 by 8 and therefore, the mutual information can now be calculated as I X Y equals H Y minus H Y given by X. So, we have H Y; H Y we have already seen is 7 by 4 minus 11 by 8 that 14 minus 11 equals 3 by 8.

(Refer Slide Time: 04:46)



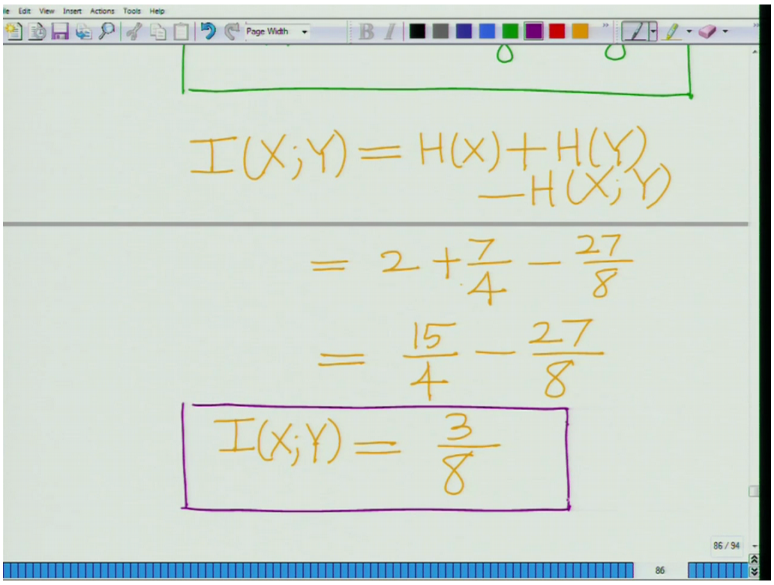
So, this is the value of your mutual information i X comma Y let us also calculated another way let us look at H what is H X given Y now H Y this is also serves as quick summary of the various review of the various property H X Y plus H X given Y equals H X comma Y implies H the conditional entropy of X given Y equals the joint entropy that is 27 by 8 minus H Y that is 7 by 4. So, that would be 27 minus 14 that is 13 divided by 8 that is your H X given Y conditional H X given Y.

(Refer Slide Time: 05:58)



Remember this is the conditional entropy of conditional entropy of X given Y conditional entropy of X given Y and again H X minus H X given Y remember H X equal 2 minus H X given Y equals 13 by. So, 16 minus 13 again 3 by 8 this is basically your i of X Y this is also your by i of X Y and yet there is a third formula remember the mutual information between X and Y that is i of X Y is also equal to H X plus H Y minus H Y comma X.

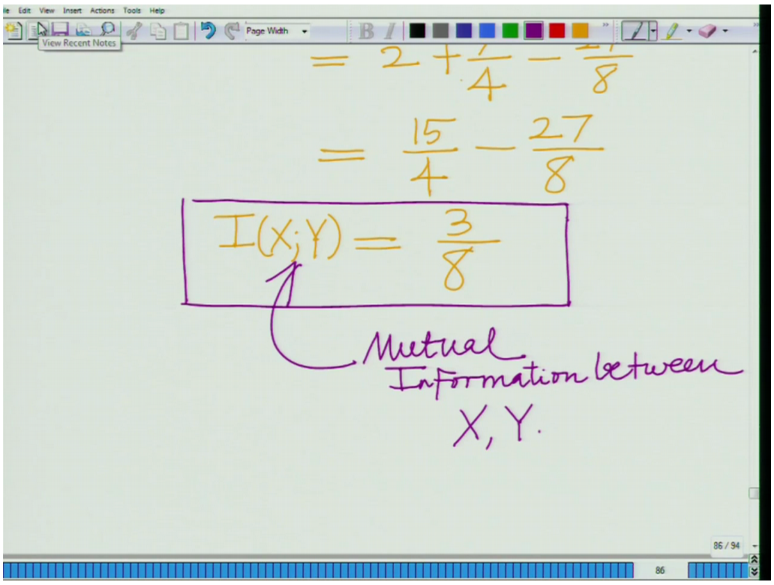
(Refer Slide Time: 06:55)



So, let us try yet another formula H X plus H Y minus H X comma Y which is equal to 2 plus s even by 4 minus 27 by 8 and if you look at this; this is basically this is 8 plus 7 that is 15 by 4. So, this is basically you are 27 by 8. So, this is basically 8 plus 7 15 divided by 4 minus 27 divided by 8 which is again 3 divided by 8.

So, basically it all comes to the same thing that is 3 divided by 8 that is the mutual information between these quantities all right. So, mutual information between X and Y is 3 divided by 8.

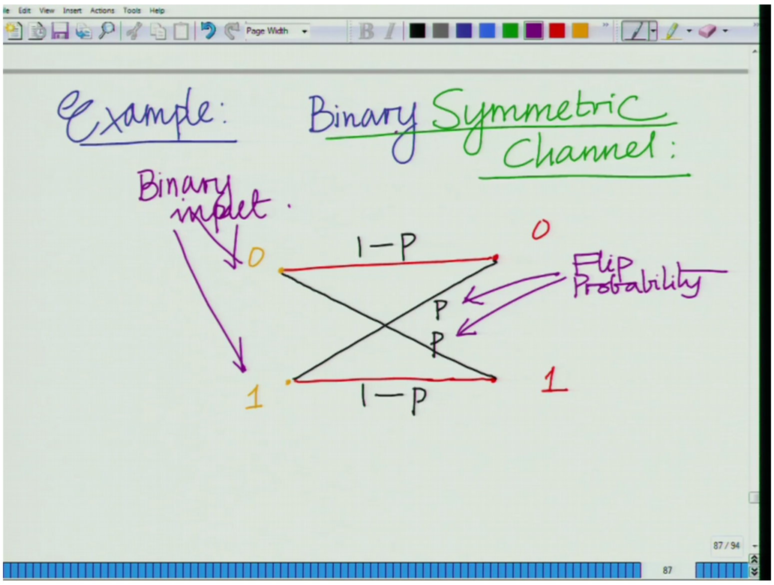
(Refer Slide Time: 08:02)



Now, let us look at a different example let us come to an example from communication system which is something that is actually very relevant because we are going to use mutual information to develop some fundamental. Fundamental principles regarding the maximum information rate that can be transmitted or the maximum rate at which information can be transmitted across all at which information maximum rate in bits per channel use at which information can be transmitted across the channel.

So, I am trying to build up towards that result. So, in that context let us look at another very relevant and meaningful result.

(Refer Slide Time: 08:58)

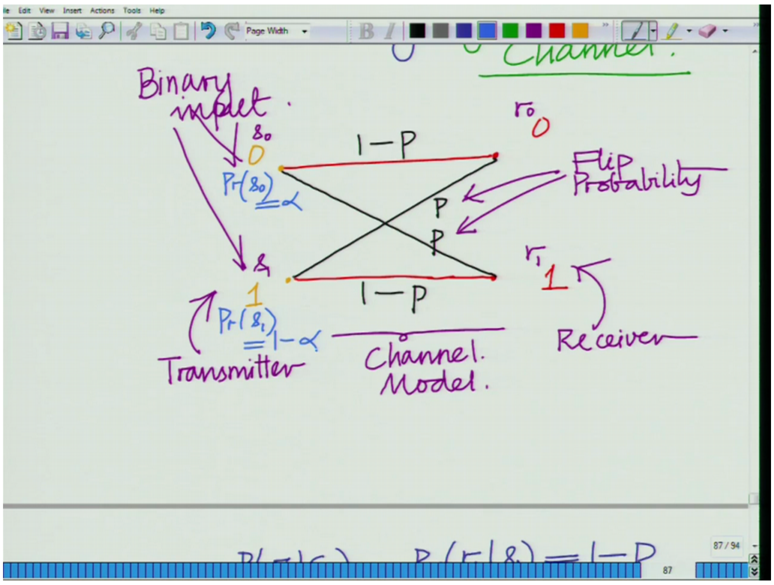


So, let us consider what is known as a binary symmetry channel some of you might already be familiar with this we are considering what is known as a binary symmetry channel in a binary symmetry channel what is have is basically let me just draw it a little bit more clearly I will have 2 let say I have a binary symbol 0 and 1 and that can be transmitted across the channel they can received either A 0 and 1.

Now, if A 0 is always received as A 1 and A 1 is always received as A 1 then there is no problem the communication is perfect, but in reality practical communication systems are far from perfect. So, A 0 there is chance of it getting flip to one A 1 there is chance of which getting flipped to 0. So, 0 getting flipped to 1 let us call the and one getting flipped to 0 let us denote this probabilities by P and therefore, naturally one getting received as 1 0 received a getting received as 0 the probabilities are. So, these are your flip probabilities

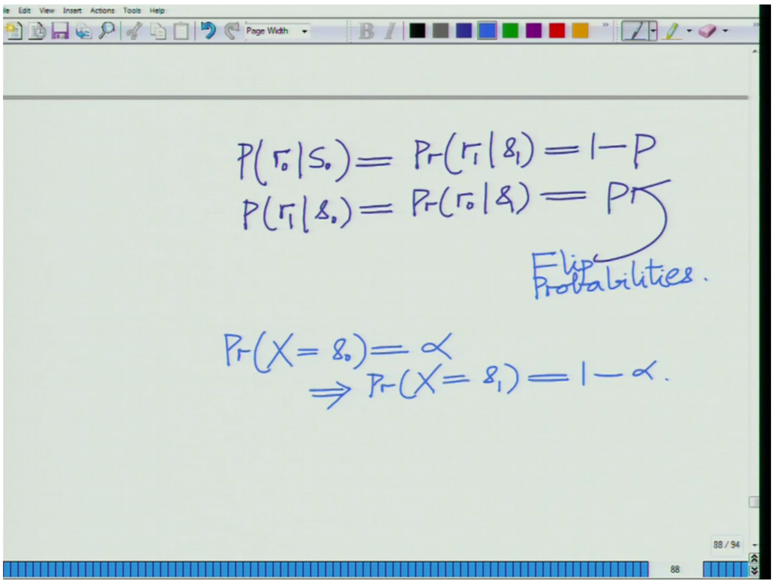
So, P is the probability with which A 0 is flipped to A 1 or A 1 is flipped to A 0 and naturally 1 minus P is a probability by which A 0 is received as A 0 and A 1 is received as A 1 across this binary symmetry channel and you call this a binary channel because its binary the inputs are binary it is 0 and 1 and symmetry because this flip probabilities are P and the direct probabilities that is the no flip where 0 is received as 0 or 1 is received as A 1 or 1 minus. So, this is symmetry with respect to 0 and 1 and therefore, this is known as a binary symmetry channel. So, this is a binary input this is a binary input.

(Refer Slide Time: 11:26)



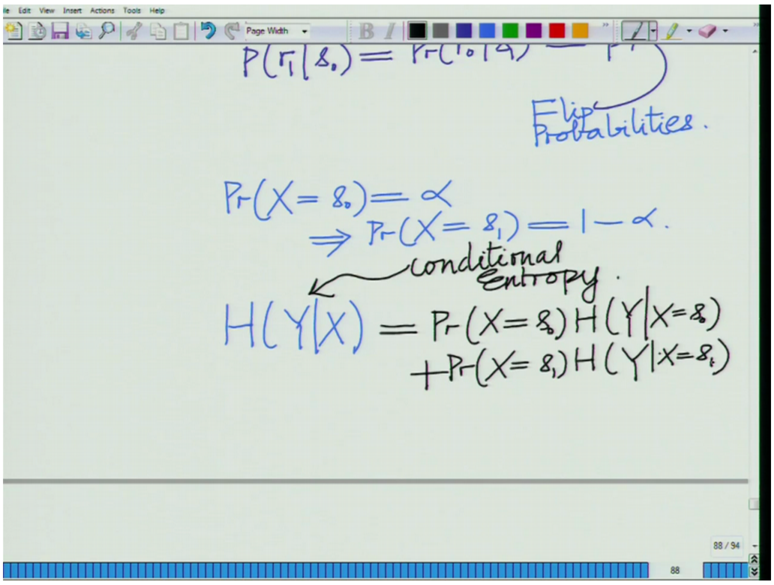
So, this is a binary input this is a flip probabilities. Now if you look at this let say this 0 we denote by 0 1 we denote by s 1 0 at the receiver we denote by r 0 and A 1 at the receiver we denote by. So, this is basically the transmitter side this is basically the; and in between is your binary symmetric. So, this binary symmetric channel is actually a channel model you can think of it as a very relevant channel model.

(Refer Slide Time: 12:06)



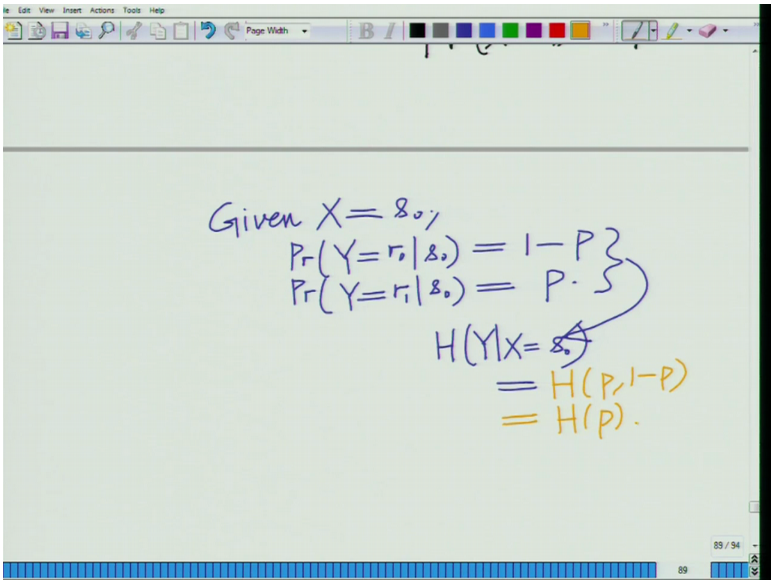
Now, if you look at this we have probability r 0 given s 0 these are the direct probabilities that is 0 is received given 0 is transmitted r 1 given s 1 that is 1 is received given 1 is transmitted equals 1 minus P and the probability r 1 given s 0 equals the probability r 0 given s 1 equals P these are the flip probabilities these are the flip probabilities. Now in addition let the probability of 0 be alpha and the probability of s 1 that is X equal to s 1 naturally will be 1 minus alpha. So, we have the probability X equal to 0 equals alpha which implies the probability X equal to s 1 equals 1 minus alpha. So, this is our binary symmetric channel probability r 0 given as 0 r 1 given as 1 are the direct probabilities 1 minus P r 0 given s 1 r 1 given s 0 are the flip probabilities which are equal to P the prior probabilities the probabilities of the symbol probabilities of symbol s 0 equals alpha probability of symbol s 1 is naturally 1 minus alpha this is a probability with which X takes either symbol s 0 or symbol s 1.

(Refer Slide Time: 13:58)



Now, what is the alpha bit right remember that is a source alpha bit. Now remember if we look at this H let us first start by computing H Y given X which is the conditional entropy the conditional entropy is equal to the probability X equal to X equal to s 0 Y given X equal to s 0 plus probability X equal to s 1 Y given X equal to s 1 now if you look at it given X equal to s 0.

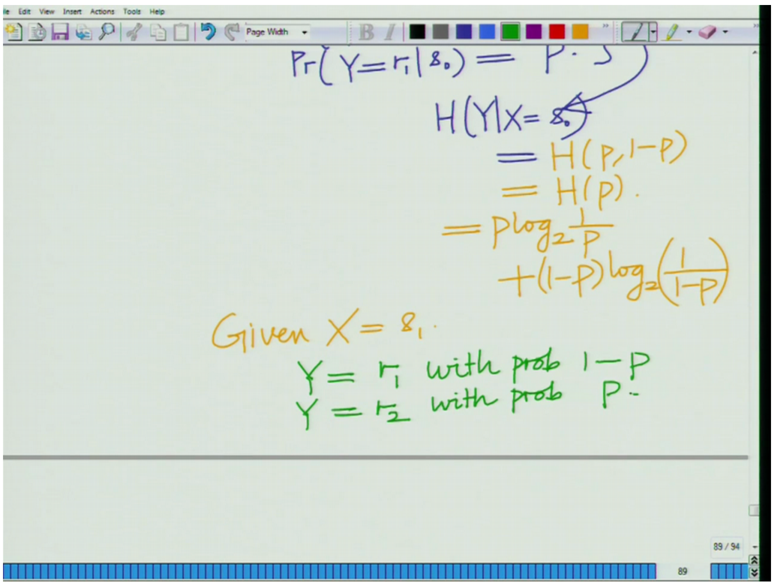
(Refer Slide Time: 14:46)



If you look at this figure given X equal to s 0 correct given X equal to s 0 Y equal to r 0 with probability 1 minus P and Y equal to r 1 with probability p

So, given X equal to s 0 probability Y equal to r 0 given s 0 equals 1 minus P similarly probability Y equals r 1 given s 0 equals P. So, therefore, entropy H of Y given X equal to s 0 corresponds to H of well these 2 probabilities P comma 1 minus P means which we simply represent by H of P.

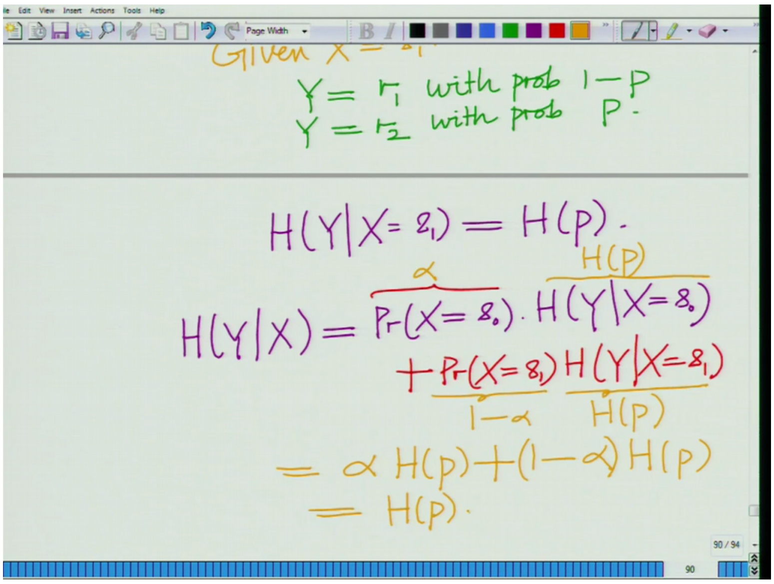
(Refer Slide Time: 15:53)



And remember since there are only 2 symbols when we say P it is immediately assumed it is assumed that there is other the other probabilities 1 minus P this is the probability of a binary source. Remember H of P denotes the probability of binary source with one of the symbols having probability P naturally the probability of the other symbols is 1 minus P and this entropy is basically nothing, but P log to the base 2 1 over P plus 1 minus P log to the base 2 1 over 1 over.

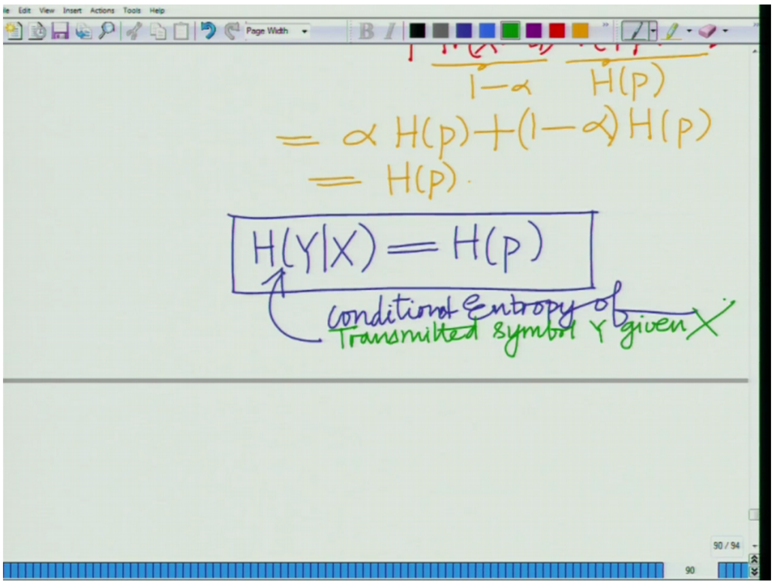
Now, similarly given X equal to s 1 given X equal to s 1 Y equal to r 1 with probability 1 minus P Y equal to r 2 with probability P.

(Refer Slide Time: 16:55)



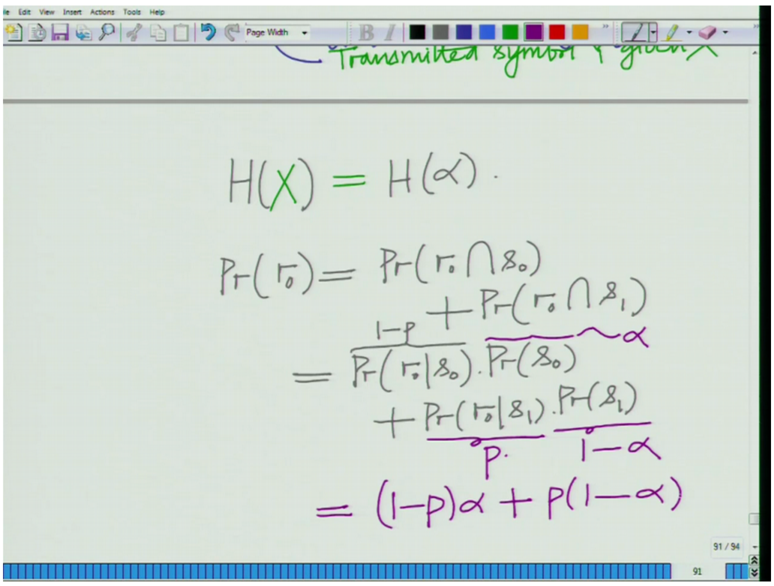
Therefore again we have H of Y given X equal to s 1 this is also equal to H of P and finally, H of Y given X equals probability X equal to s 0 into H of Y given X equal to s 0 plus probability X equal to s 1 into H of Y given X equal to s 1 this is remember alpha this is H of P this is 1 minus alpha H of Y given X equal to s 1 this is also H of P. So, this is equal to alpha times H of P plus 1 minus alpha times H of P which is equal to H of P.

(Refer Slide Time: 18:06)



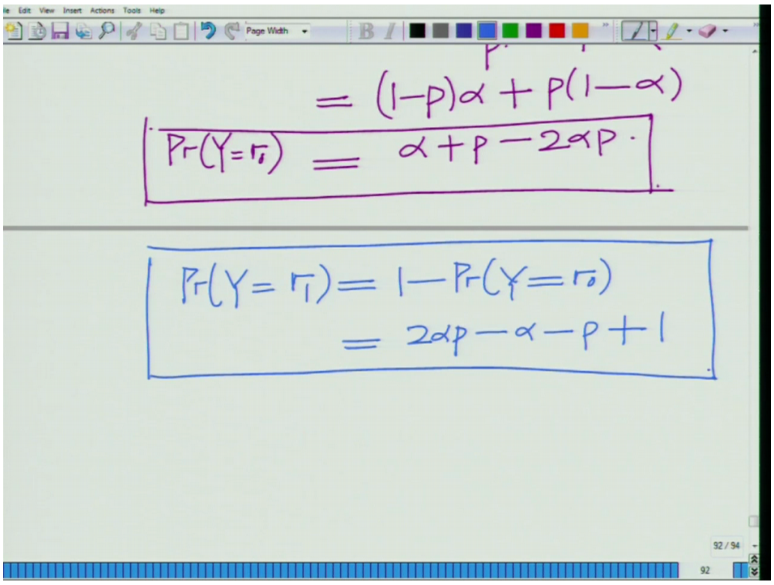
So, we have derived an important result here which is that H of Y given X equals H of P H of Y given X equals H of p. So, you have derived the conditional information the conditional entropy of Y given X that is the output Y of this channel. Remember now we are taking a practical example of a binary symmetric channel. So, this is a conditional entropy of the output Y given the transmitted symbol X conditional entropy is good to keep this in mind transmitted symbol Y given X.

(Refer Slide Time: 19:09)



Now, let us come to H of Y to compute the mutual information we need H of Y because the mutual information is H of Y minus H of Y given X of course, we know H of X if you look at this we know what is H of X H of X is basically equal to that as symbols with probability alpha and 1 minus alpha. This we know now if you look at probability of r 0 that is equal to probability r 0 means Y equal to r 0 r 0 intersection is 0 from the total probability rule this is a probability r 0 intersection s 0 plus probability r 0 intersection s 1 which is probability r 0 given s 0 into probability s 0 plus probability r 0 given s 1 into probability s 1 we know this; this is the direct probability 1 minus P we know this; this is probability of s 0 which is alpha this is the flip probability probability r 0 minus 1 that is P probability of s 1 that is 1 minus alpha.

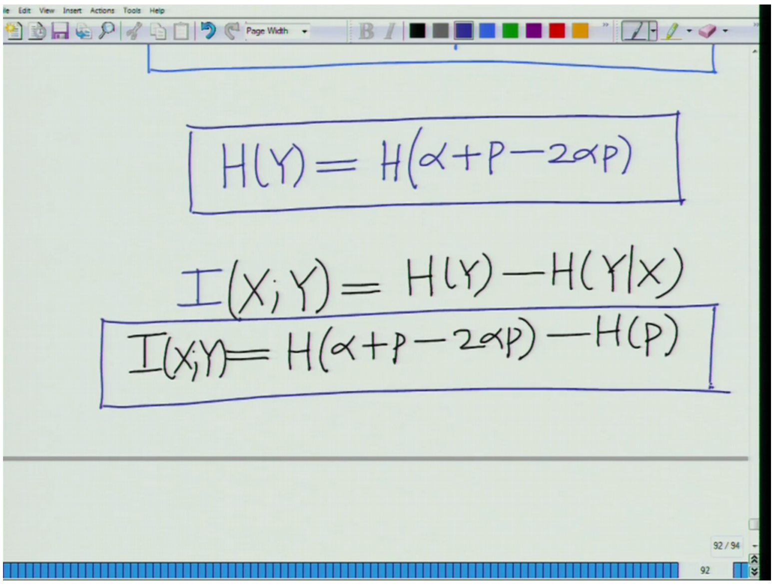
(Refer Slide Time: 20:37)



So, therefore, this is 1 minus P times alpha plus P into 1 minus alpha which is equal to well alpha plus P minus twice alpha P that is your probability of r 0 or basically let say probability of Y equal to r 0 to be more explicit leaves not confusion and therefore, probability of Y equal to r 1 is naturally 1 minus this probability of r 1 is minus probability of r 0. So, probability Y equals r 1 equals 1 minus probability Y equal to r 0 that is equal to a basically that will be 2 alpha P minus alpha minus P plus 1.

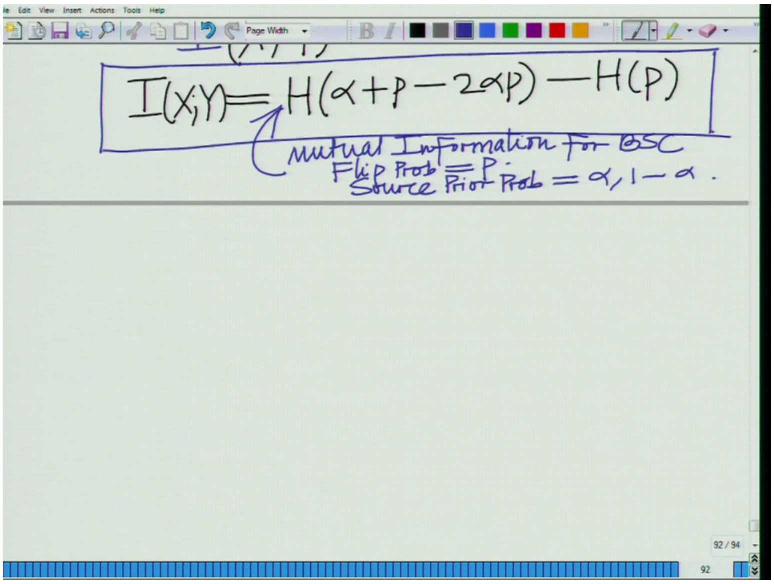
So, this is your probability Y equal to r 1 and therefore, entropy and therefore, the entropy H Y.

(Refer Slide Time: 21:51)



Remember H Y is also binary source with 0 and 1 is equal to nothing, but H of alpha plus P minus 2 alpha P remember this is a binary source probability of r 0 as alpha plus P minus 2 alpha P probability of the other symbol is actually 1 minus this. So, H of Y is basically simply H of alphas plus P minus twice alpha P and therefore, the mutual information. Now we are able to find the mutual information for the binary symmetric channel this is equal to H Y minus H Y given X that is equal to H of alpha plus P minus twice alpha P minus H of P this is your mutual information.

(Refer Slide Time: 22:58)



So, this is the mutual information for the binary symmetric channel flip probability equals P and source or transmit symbol probability alpha comma 1 minus alpha correct. So, here probability is alpha 1 minus alpha that is probabilities equal to r s 0 the s 0 and s 1 occur with probability of alpha at 1 minus alpha and the channel is a binary symmetric channel with flip probability P. And for that we have derived the mutual information and the mutual information is H of alpha plus P minus 2 alpha P minus H of y; which is H of Y minus H of Y given X.

And similarly you can also calculated H of X minus H of X given Y and it will come out to the same value you have to you can verify that will come out to the same value. So, in this module what we have seen is we have seen the mutual of course, we have defined the mutual information the previous module, but we have completed that. Now with that several relevant examples one of course, is the simple example based on a set of joint probabilities of course, is synthetic examples based on set of joint probabilities that we have been considering to illustrate the various quantities information theory quantities alright.

The other examples is a very relevant and a practical example of a binary symmetric channel in which we have 2 transmit symbols that is 0 and 1 right in which the transmit that is the alphabet the transmit symbol alphabet source alphabet comprising of 2 symbols 0 and 1 which can be other received as 0 or 1 each the flip probability is P the direct probability is 1 minus P and for this channel we have characterize the mutual information. And now we are going realize that in the next module in the subsequent modules we are going to see the importance of this mutual information that is how the mutual information relates to the information rate that can be transmitted or the information rate that can be conveyed across this channel or for that metal any general channel which is termed as the capacity of the channel which is a very fundamental result which is one of the many fundamental results in information theory. So, that we are going to deal in the subsequent modules.

Thank you very much.