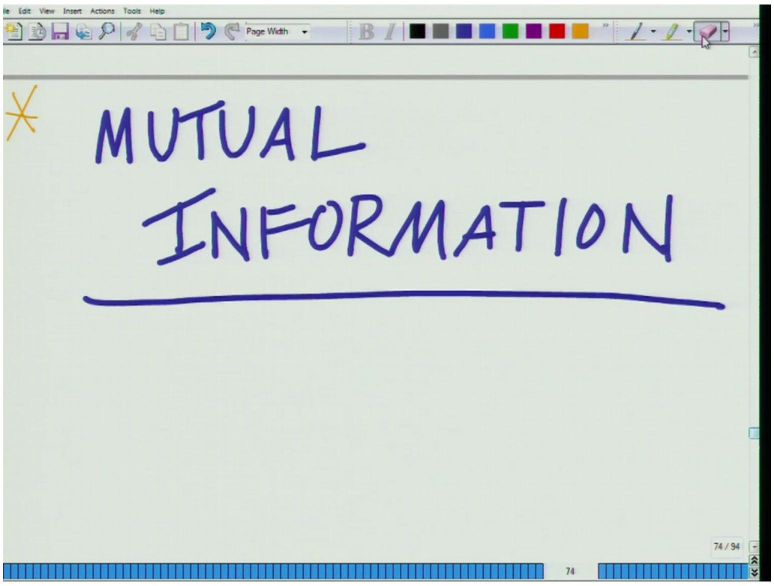
Principles of Communication Systems - Part II
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture - 33
Mutual Information, Diagrammatic Representation, Properties of Mutual Information

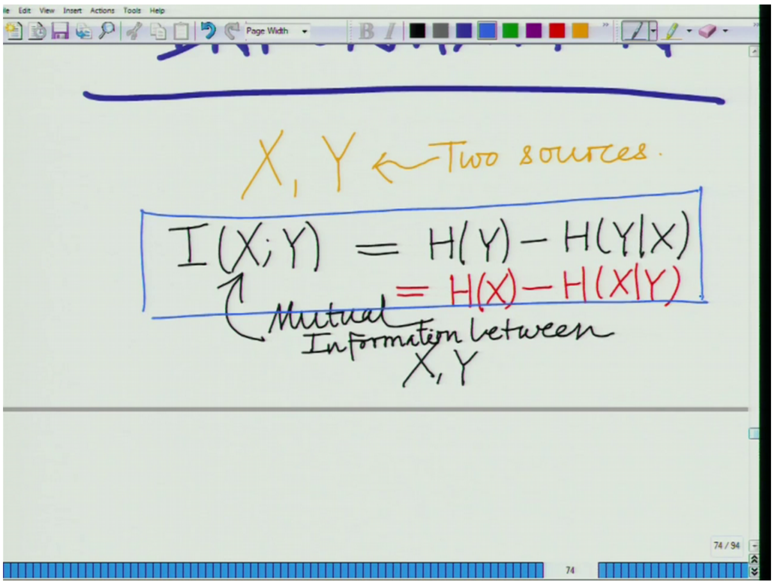
Hello welcome to another module in this massive open online course. Today let us look at another new concept information theory that is a concept of mutual information ok.

(Refer Slide Time: 00:26)



So, let us look at this new concept. So, this is mutual information. And as the name implies it is the mutual information it is, what is mutual when you talk about the term mutual between 2 things right. So, it characterizes the interpreter, characterizes this relationship between the information of these of 2 sources on X. So, let us consider 2 sources let us X and Y.

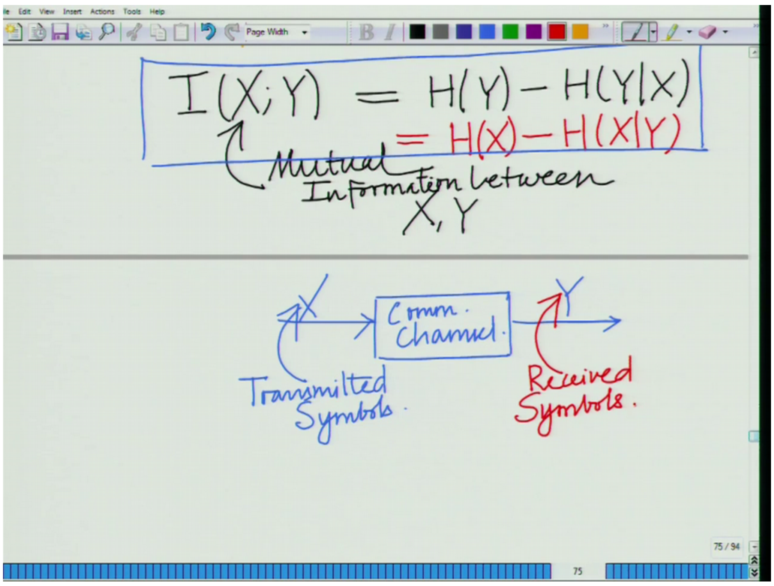
(Refer Slide Time: 01:15).



And this is one of the fundamental quantities that can be used to analyze the performance of a communication system alright.

So, I have X and Y, which are 2 sources or have alphabet drawn from 2 sources. Then the mutual information between X and Y that is denoted by i X semicolon Y, correct? This is the mutual information, information between X and Y that is equal to H, defined as H Y minus H Y given X, which is also equal to H X minus H X given Y. This is the definition of the quantity, this is the definition of this term mutual information between X.

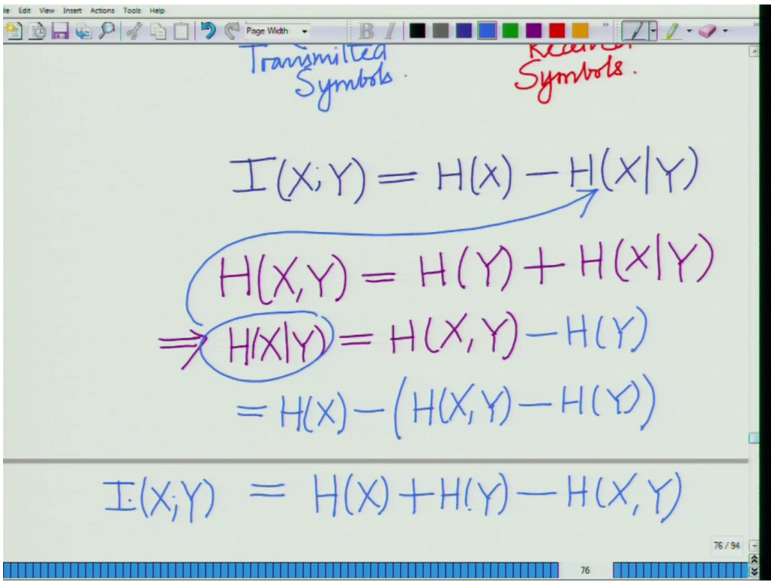
(Refer Slide Time: 02:31)



For instance you can have as we are considering our aim is to characterize the fundamental limit on the performance of a communication systems. So, I can X which are the transmitted symbols Y, which are the so this can be for instance a communication channel, these are the transmitted symbols.

This can be the received symbols, transmitted symbols received symbols. And basically what we have is the mutual information between X and Y is equal to H of X minus H of X given Y.

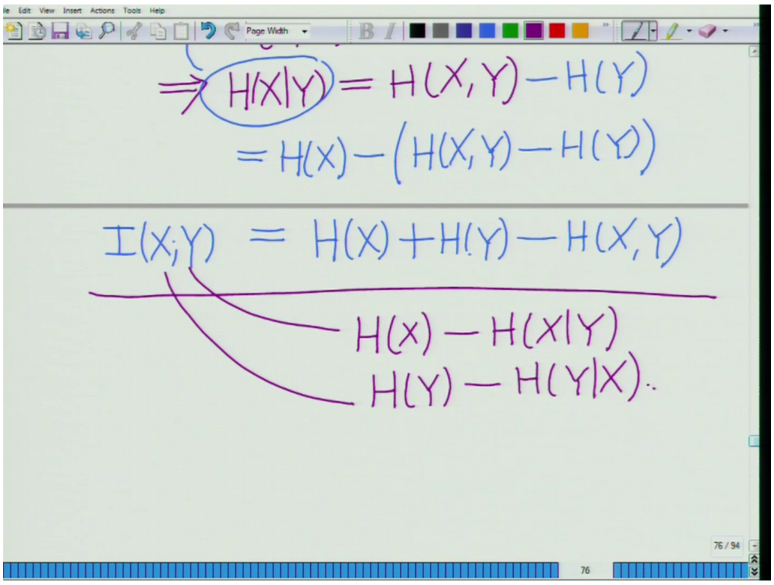
(Refer Slide Time: 03:17)



Now, look at this we have, we know from our joint entropy we know that H of X comma Y equals H of Y plus H of X given Y which employs that H of X given Y equals the joint entropy H of X comma Y minus the entropy H of Y; right let me try this again. So, this employs that H of X given Y equals the joint entropy is of X comma Y minus the entropy H of Y. I am going to substitute this quantity in here and that gives me H of X minus H of X given Y which is nothing but H of X comma Y minus H of Y, which is equal to H of X plus H of Y minus H of X comma Y.

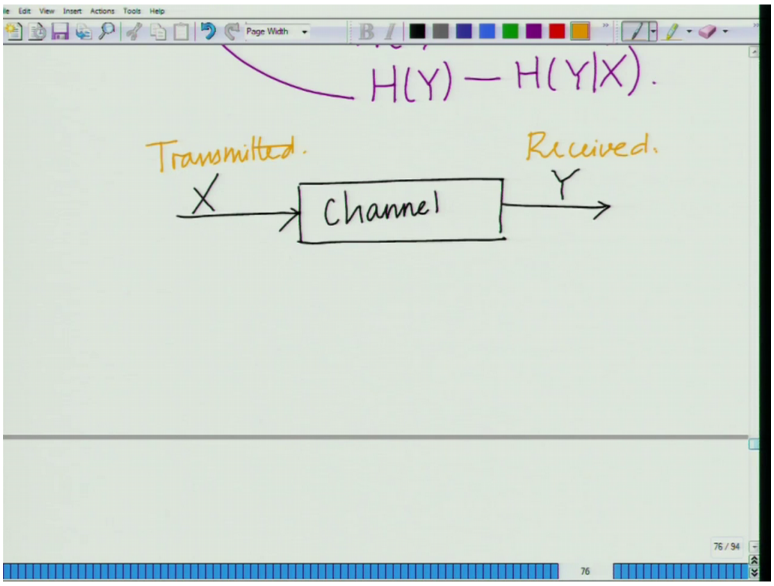
So, therefore, this is the equivalent definition for the mutual information alright. So, I have X Y is H of X plus H of Y minus H of minus, that is H of X plus H of Y minus H X comma Y ok.

(Refer Slide Time: 05:27)



So, these are all equivalent definitions also H of X minus H of X given Y this is also H of Y minus H of Y given X.

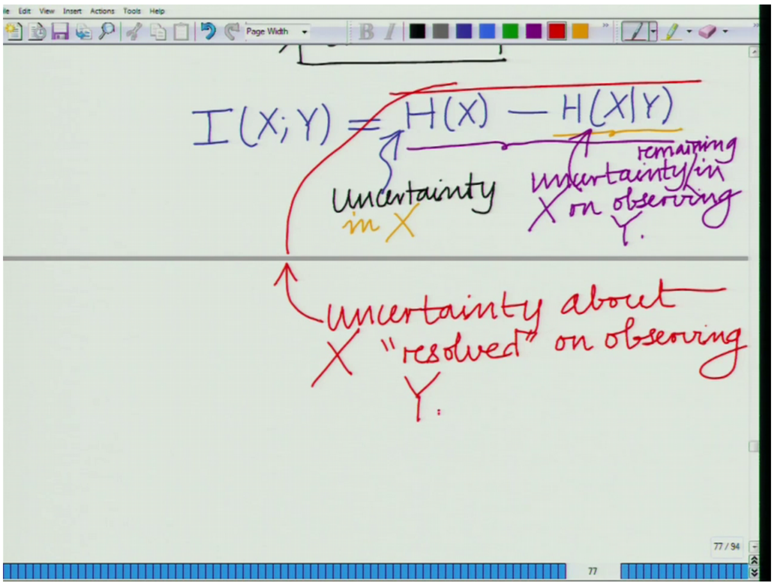
(Refer Slide Time: 05:49).



And now if you can look at this definition. Let us look at this So, let us look at this again in the context of communication. Now if you look at this in the context of communication. So, let us say this is our communication channel this is Y X is what is transmitted. Now our aim and communication is so, this is transmitted. Now if you look at a communications scenario our aim is at the receiver receive Y receive the output Y output symbols Y. And infer or make decisions regarding the transmitter symbol X, alright? From Y we have to infer what is what, what are the symbols X that have been transmitter. So, that is the purpose of the communication, look at the receiver look at the output of the receiver and infer about the inputs or the transmitted symbols or the inputs to the communication channel or the transmitted symbols alright.

So, we have to infer what has been transmitted by the transmitter. Now, so therefore, if I look now at the mutual information between X comma Y, the mutual information between X comma Y is basically information in X H X minus H X given Y.

(Refer Slide Time: 07:08).

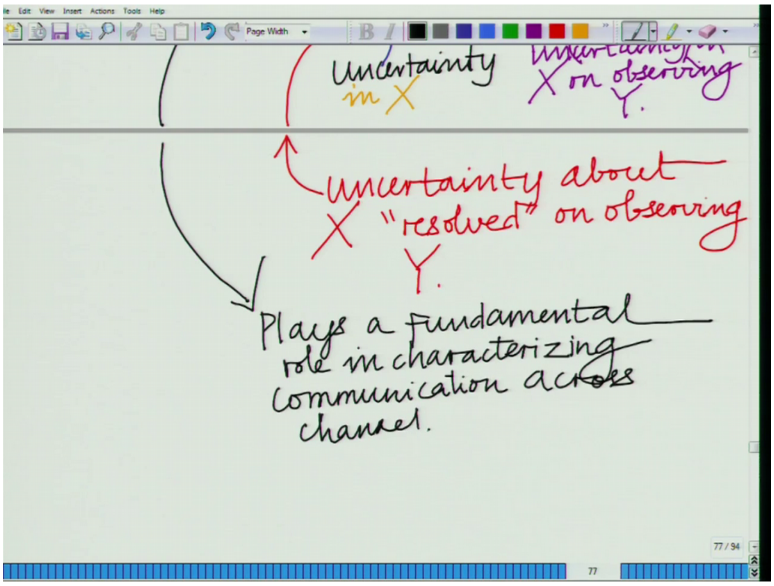


Now naturally If you look at the information definition in terms of uncertainty H X, H X is the uncertainty in X the uncertainty, uncertainty in the transmitter symbols X. H X given Y is the conditional uncertainty or basically uncertainty remaining in X having observed by. So basically, H X is the uncertainty in X, H X given Y that is having absurd Y, what is uncertainty that is remaining in X? Right.

Now, naturally if Y conveys a lot of information about X or observing Y, if you are able to resolve X or if you are able to resolve X to good degree or if we able to estimate transmitted symbols X to a good degree of accuracy. Then the uncertainty remaining in X having observed Y that is H of X given Y must be low. Now, therefore, H of X minus H of X given Y if you look at this H of X is the uncertainty in X, H of X given Y is the uncertainty in X or observing Y that is uncertainty remaining in X, you can say uncertainty remaining in X, remaining in X on observing. So, H of X minus H of X given Y, this whole quantity is the uncertainty about X resolved after observing Y, after X that is resolved. Uncertainty about X resolved on observing Y, H of X is the uncertainty in X H of X is the uncertainty remaining after observing Y we expect this quantity to be less.

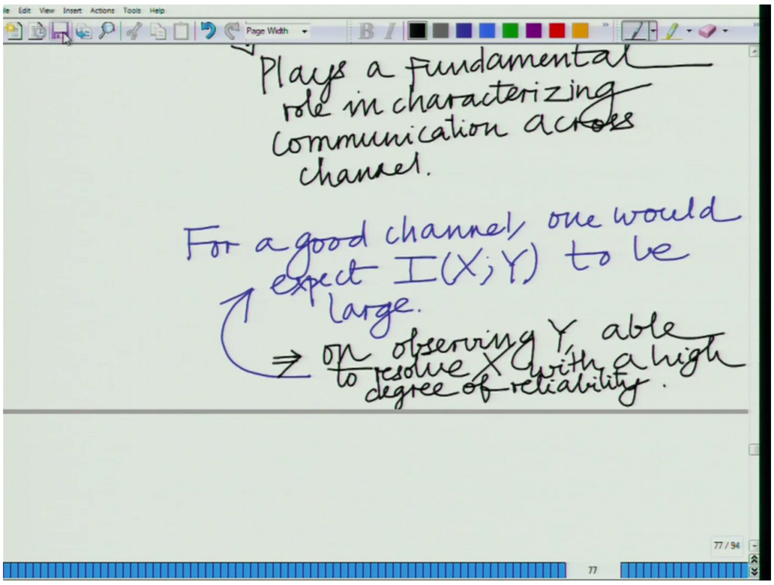
So, H of X minus H of X given Y is the uncertainty that is resolved uncertainty about X that is resolved on observing Y. Or in another sense this is the information right, this information that is basically information, that is that has been transported across that this is this is basically the information that has been conveyed across the channel alright. So, basically H X is the information X the uncertainty in X, H X given Y is the uncertainty remaining. So, uncertainty resolved after observing Y is H X minus H X given Y which is basically you can think of this as the information that has been conveyed by transmission of the symbols through the channel alright. So, that is the information that is the mutual information i X semicolon i X Y is uncertainty about X that is resolved on observing Y ok.

(Refer Slide Time: 10:47)



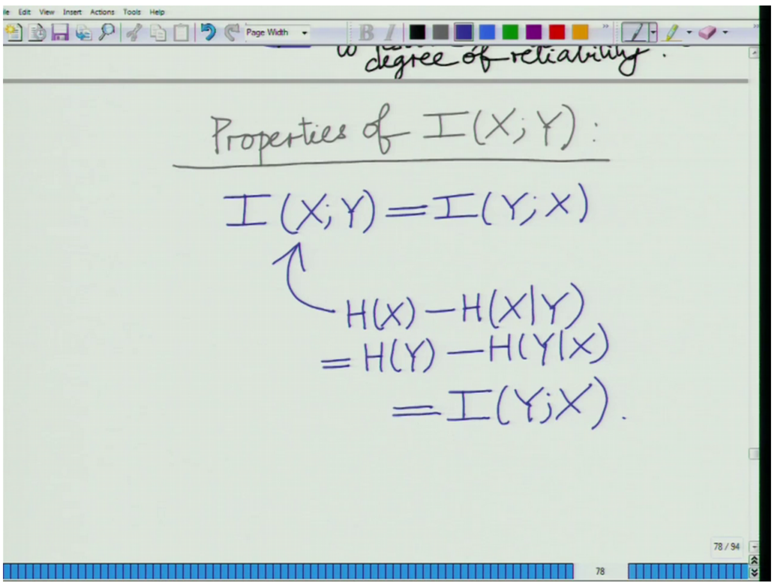
And this therefore, has a very fundamental role to play in communication the mutual information therefore, plays as we shall see later plays a fundamental role in characterizing the performance of a, plays a fundamental role in characterizing communication across the channel.

(Refer Slide Time: 11:31).



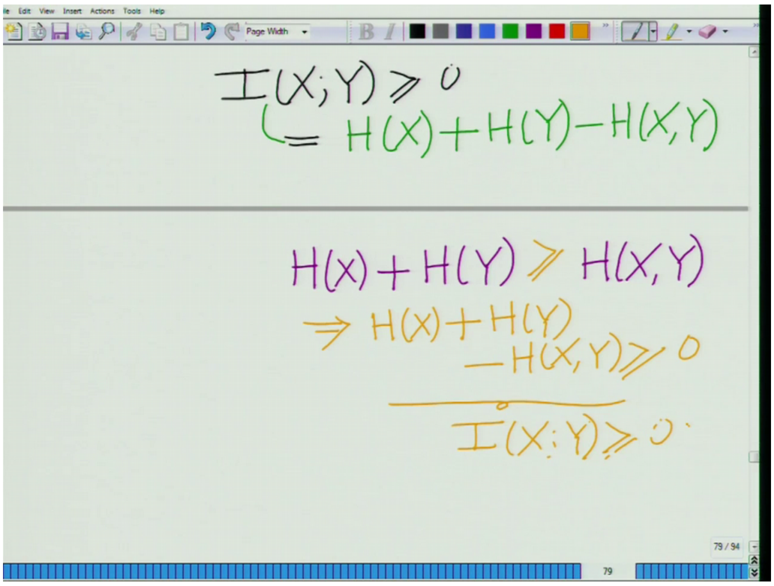
So, for a good channel naturally. One would expect i of X Y to be large. i of X Y to be large that is, on observing Y one should be able to resolve X to a with a high degree of accuracy. With a high degree of accuracy or with a high degree of reliability, with a high degree of reliability, that is same. So, basically mutual information as we see has a said has a key role to play in characterizing the efficiency of communication across a channel, alright? Will is we will discuss this further look at the maximum will look at the information rate across a channel alright.

(Refer Slide Time: 13:06).



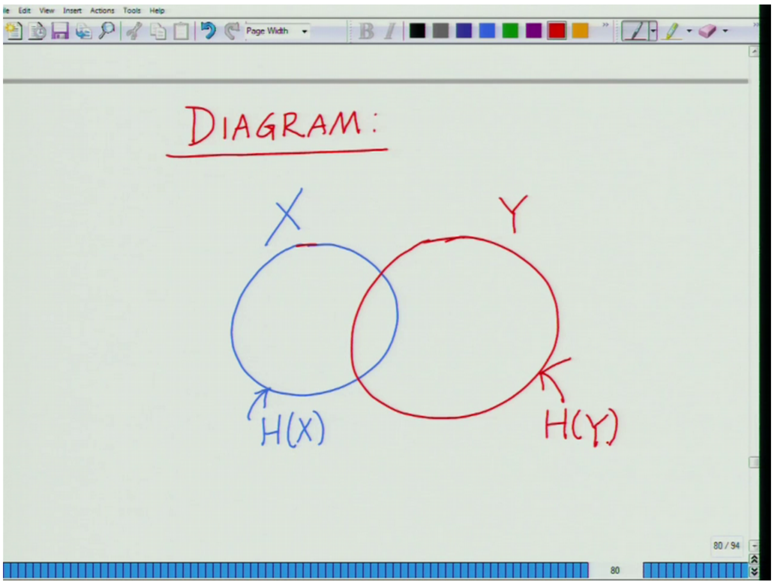
Now let us look at some properties of this mutual information. Now the properties of mutual information are that i X comma Y equals i Y comma X because we have seen this is nothing but H of X minus H of X given Y, which is equal to H of Y minus H of Y given X, which is equal to i of Y comma X no surprises there. So, it is very simple mutual information between X and Y is same as the mutual information between Y and X.

(Refer Slide Time: 14:07).



Now, also it should be clear to you mutual information between X and Y is greater than or equal to 0. Because the mutual information remember, can be written as H X plus H Y minus H X comma Y, and if you remember we have seen that H X plus H Y is less than or equal to H X Y. We have already seen this which implies H X plus H Y minus H, I am sorry H X, X Y is greater than or equal to H X Y. H X Y, H Y implies H X plus H Y minus H X H, X Y is greater than or equal to 0 and this nothing but the mutual information, mutual information between X comma Y is greater than or equal to greater than or equal to 0. And we can have a pictorial representation of the mutual information this is often convenient.

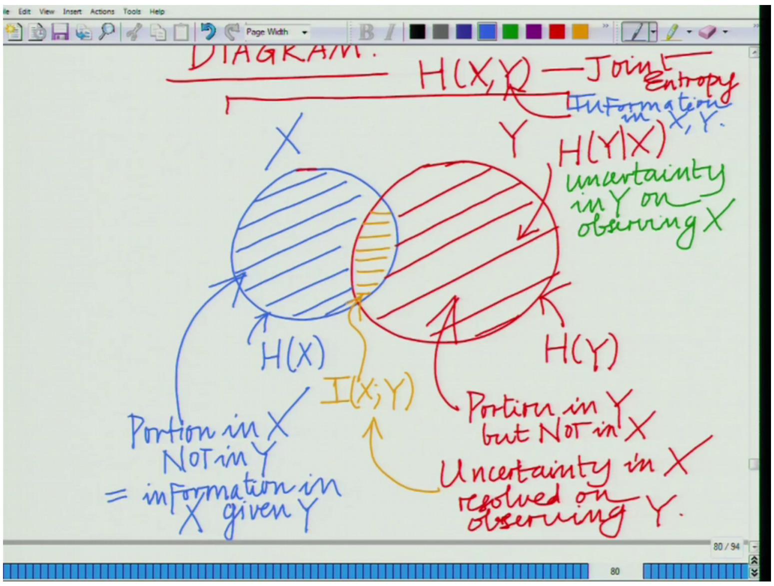
(Refer Slide Time: 15:36)



So, let us draw a diagram in fact, not just mutual information, but the various measures of information that we have seen.

So, let us say we have 2 sources one is X the other is Y. So, this is your source X this is Y. Now if you look at this let me shade this with blue this whole thing represents the information H X and I shade in red this whole thing represents, or just to keep it simple let me not shade these region, this whole thing represents H X this whole thing represents H Y information X information Y.

(Refer Slide Time: 16:53)

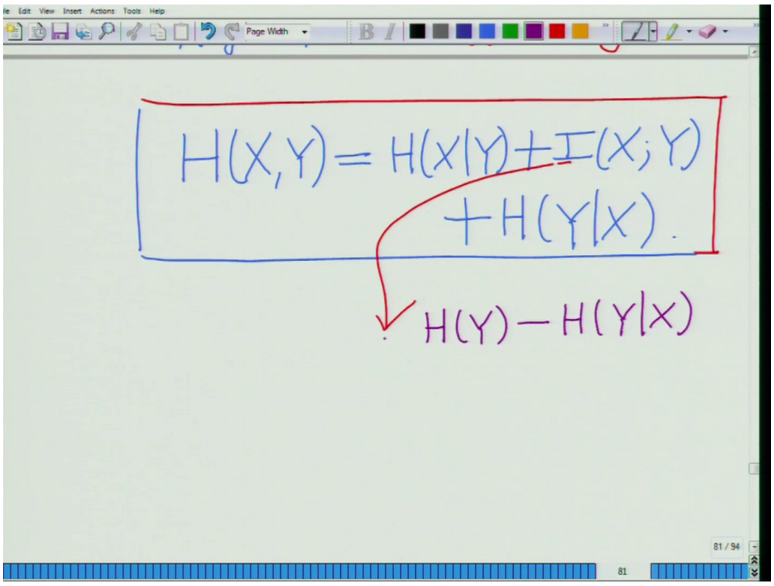


Now, if you look at these components individually. If you look at this component this is the uncertainty remaining this is uncertainty remaining in X. If you look at this shaded region that is in y, but not in X this is uncertainty remaining in Y on observing X.

So, this is uncertainty in Y on observing X or conditioned X alright. So, we have information in X information in Y. Portion in Y not in X is the uncertainty or the information in Y remaining in Y. So, this is basically, this is basically precisely information remaining in X that is portion let me write it this portion in y, but not in X. Similarly the portion in x, but not in Y is the information in X on observing Y or uncertainty. So, portion in X information or X, information in X given Y or uncertainty X on observing Y. And if you look at this area of overlap that is portions in both X and Y this is basically the mutual information that is information resolved information about X resolved on Y resolving and observing X or uncertainty in X resolved on observing Y.

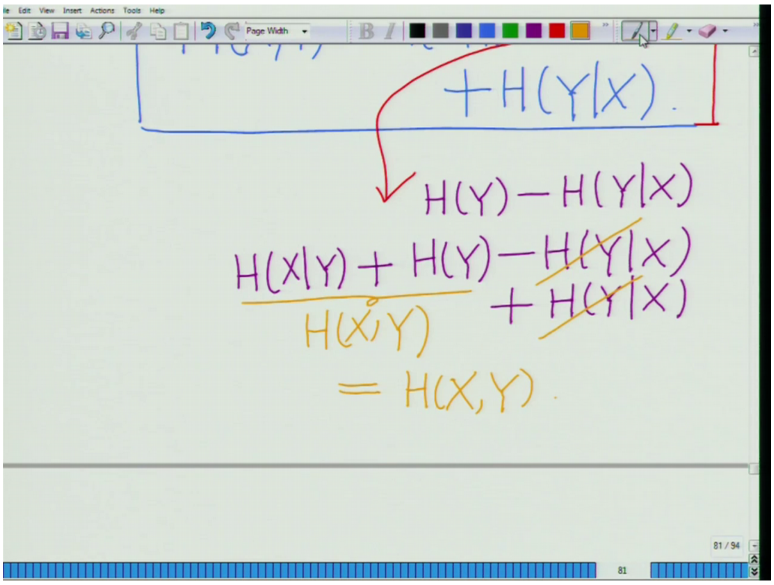
So, we can say this is the uncertainty in X resolved on observing Y. And if you look at this total quantity, now this is the joint entropy information in X and Y this is joint entropy to get both this together. This is information in X comma Y. This is the joint entropy that is information in X and Y. And if you observe this figure, now you can see the joint entropy H X Y is a sum of 3 mutually disjoint components, H this is H X given Y mutual information between X and Y and H Y given X.

(Refer Slide Time: 20:33).



So, the now you can write you can see from the figure H X comma Y H X given Y plus the mutual information i X Y plus the information H i Y H Y given X. And this you can see from the figure and this you can verify also for instance, i X given Y this is equal to H Y minus H Y given X.

(Refer Slide Time: 21:18)



So, substituting we have H X given Y plus i X Y which is mutual information between X and Y H Y minus H Y given X plus H Y given X, these cancel these cancel and now if you look at this H X given Y plus H Y. This is nothing but the joint entropy H X comma Y.

So, we can derive such useful results using the mutual information. So, this is the property and you can, and you can look this these are 3 disjoint components H X given Y the mutual information between X Y, and H Y given X that is the uncertainty in Y on observing X. And therefore, that brings us end of this module, basically what you have seen in this module is we have defined a new measure of which, I would rather key or a rather critical measure to understand the subsequent theory that we are going to develop to characterize the fundamental rate of information transfer that is fundamental rate at which information can be transmitted across the channel. This is termed as the mutual information.

It is defined as H X between 2 sources it is defined as H X minus H X given Y or the H Y minus H Y the mutual information is already always greater than equal to 0 and finally, we have seen the relevance of mutual information that is mutual information between X and Y is the information about the uncertainty about X that is resolved on observing Y. Or you can also think of it as the uncertainty about Y that is resolved on observing X, it is symmetric mutual information between X Y is the same as the mutual information between Y and X. And it has a fundamental role to play in characterizing the rate or how much information arise at what rate can information be transmitted across the channel.

And further we have seen a pictorial way to sort of comprehensively represent the various quantities that we have seen so far the entropies the joint entropy the conditional entropy and now also the mutual information between 2 sources alright.

So, we will stop here and continue with the other aspects in the subsequent modules.