Mutual Information

Mutual information is one of many quantities that measures how much one random variables gives about another. It is a dimensionless quantity. Mutual Informaiton can be thought of as the reduction in uncertainty about one random variable given knowledge of another.

- High mutual information indicates a large reduction in uncertainty,
- low mutual information indicates a small reduction,
- zero mutual information between two random variables means the variables are independent.

Efficient communication systems have high mutual information.

Mutual Information

Joint Entropies:

Using the input probabilities $P(x_i)$, output probabilities $P(y_i)$, transition probabilities $P(y_i|x_i)$, and joint probabilities $P(x_i, y_j)$, we can define the following various entropy functions for a channel with m inputs and n outputs:

- $H(X) = -\sum_{i=1}^{m} P(x_i) \log_2 P(x_i)$
- $H(Y) = -\sum_{j=1}^{n} P(y_j) \log_2 P(y_j)$
- $H(X,Y) = -\sum_{j=1}^{n} \sum_{i=1}^{m} P(x_i, y_j) \log_2 P(x_i, y_j)$

Mutual Information: Joint Entropy

These entropies can be interpreted as follows:

- $H(X)$ is the average uncertainty of the channel input, and $H(Y)$ is the average uncertainty of the channel output.
- The joint entropy $H(X,Y)$ is the average uncertainty of the communication channel as a whole.

Mutual Information: Conditional Entropy

- The conditional entropy $H(X|Y)$ is a measure of the average uncertainty remaining about the channel input after the channel output has been observed.
- This is sometimes called the equivocation of X with respect to Y.
- The conditional entropy $H(Y|X)$ is the average uncertainty of the channel output given that X was transmitted.

- $H(X|Y) = -\sum_{j=1}^{n}\sum_{i=1}^{m} P(x_i, y_j)\log_2 P(x_i|y_j)$
- $H(Y|X) = -\sum_{j=1}^{n}\sum_{i=1}^{m} P(x_i, y_j)\log_2 P(y_j|x_i)$

Mutual Information : Useful Identities

Two useful relationships among the types of entropies are

- $H(X,Y) = H(X|Y) + H(Y)$
- $H(X,Y) = H(Y|X) + H(X)$

(Remark : compare to identities in probability theory)

Mutual Information : Useful Identities

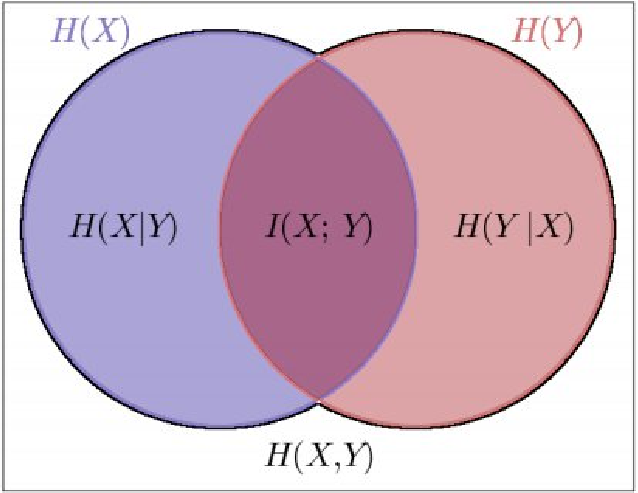

Figure:

Mutual Information

The mutual information $I(X;Y)$ of a channel is defined by

$$I(X;Y) = H(X) - H(X|Y) \text{ (b/symbol)}$$

Alternatively we can define it as either

$$I(X;Y) = H(Y) - H(Y|X) \text{ (b/symbol)}$$

or as

$$I(X;Y) = H(Y) + H(Y) - H(X,Y) \text{ (b/symbol)}$$

Remark: The mutual information is the reduction of entropy of X when Y is known.