Communications Theory
Lab Session 3: Communications Theory
Academic Year 2023/2024

Goals

In this lab, the student will learn to

- It is recommended to read the statement and the theoretical content within it before starting to work on the practice.
- The lab session must be carried out in pairs.
- Each pair must complete the questionnaire.
- The deadline is after the lab session. Late submissions will imply a penalty.

Rules and deadline

- The practice can be carried out in pairs.
- It is recommended to read the statement and the theoretical content within it before starting to work on the practice.
- At the end of the lab session, the student must hand in the questionnaire filled out. Later on, she/he will upload to *Aula Global* the source code on which she/he relied to answer the questions.

1 Fundamental theoretical concepts

This lab session focuses on the estimation of quantitative information metrics from realizations of random variables, and the capacity of a related channel. In this section we briefly review the theoretical concepts required to do the lab.

1.1 Quantitative measures of information

In this section, several quantitative information metrics for discrete random variables are presented. Regarding notation, we consider two random variables X and Y whose alphabets have, respectively, M_X and M_Y elements,

$$\mathcal{A}_X = \{x_0, x_1, \cdots, x_{M_X-1}\},\ \mathcal{A}_Y = \{y_0, y_1, \cdots, y_{M_Y-1}\},$$

and probability distributions (probability density/mass functions)

$$p_X(x_i) \text{ y } p_Y(y_j), \text{ for } i \in \{0, 1, \cdots, M_X - 1\},\ j \in \{0, 1, \cdots, M_Y - 1\}.$$

Finally, the joint probability of the two random variables is denoted $p_{X,Y}(x_i, y_j)$.

Next, the information metrics are introduced.

1.1.1 Entropy of a random variable

The entropy of a random variable is the average of the *autoinformation* of every element in the alphabet thereof

$$H(X) = -\sum_{i=0}^{M_X-1} p_X(x_i)\ \log p_X(x_i) = \sum_{i=0}^{M_X-1} p_X(x_i)\ \log\left(\frac{1}{p_X(x_i)}\right).$$

Recall that units depends on the base of the logarithms. In this lab, logarithms to base 2 are used, so that units are bits/symbol.

Also, it is useful to keep in mind that in this and subsequent definitions, we have $0 \times \log(0) = 0$.

1.1.2 Joint entropy of two random variables

The joint entropy of two random variables is the natural extension of the concept of entropy for a single random variable to the case of two random variables where the average is now over every pair of values both variables can take

$$H(X,Y) = -\sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{XY}(x_i,y_j)\ \log\left(p_{XY}(x_i,y_j)\right) = \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{XY}(x_i,y_j)\ \log\left(\frac{1}{p_{XY}(x_i,y_j)}\right).$$

1.1.3 Conditional entropy

Conditional entropy is defined as the average of the entropy of a random variable conditional on every value of the other random variable

$$H(X|Y) = \sum_{j=0}^{M_Y-1} p_Y(y_j)\ H(X|Y=y_j).$$

Straightforward algebraic manipulations allow to express this entropy using the conditional probabilities, $p_{X|Y}(x_i|y_j)$, as

$$H(X|Y) = \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{XY}(x_i,y_j)\ \log\frac{1}{p_{X|Y}(x_i|y_j)}.$$

It is good to recall the connection between joint, conditional and marginal probabilities stemming from Bayes rule,

$$p_{X|Y}(x_i|y_j)\ p_Y(y_j) = p_{XY}(x_i,y_j).$$

1.1.4 Mutual information

Mutual information is defined as

$$I(X,Y) = \sum_{i=0}^{M_X-1}\sum_{j=0}^{M_Y-1} p_{XY}(x_i,y_j)\ \log\frac{p_{XY}(x_i,y_j)}{p_X(x_i)\ p_Y(y_j)}.$$

It is related to the entropy through several formulas that are usually represented by means of a Venn diagram such as the one shown in Figure 1. These formulas are

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

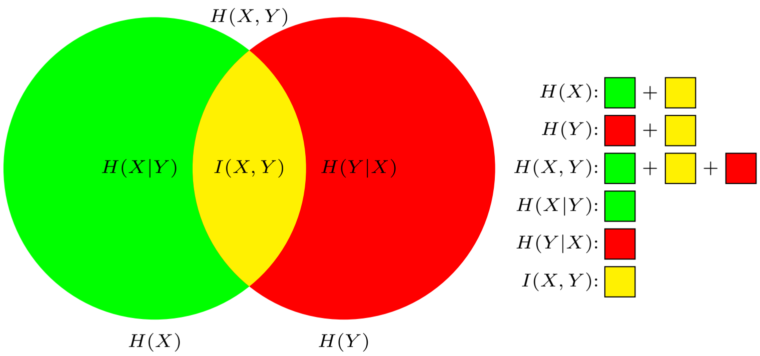

Figure 1: Connections between different information theory-related metrics through a colour code.

1.2 Probabilistic channel models for discrete channels

The probabilistic model for a discrete channel in which the inputs is modeled as a random variable X, and the output as as another random variable Y, is given by the so-called *transition probabilities*, which are simply the conditional probabilities of the outputs given the input, $p_{Y|X}(y_j|x_i)$,, for every possible input-output combination.

In representing these probabilities, they are often gather in a *channel matrix*, which is a matrix with M_X rows and M_Y columns which arranges the transition probabilities as follows

$$\mathbf{P}(Y|X) = \begin{bmatrix} p_{Y|X}(y_0|x_0) & p_{Y|X}(y_1|x_0) & \cdots & p_{Y|X}(y_{M_Y-1}|x_0) \\ p_{Y|X}(y_0|x_1) & p_{Y|X}(y_1|x_1) & \cdots & p_{Y|X}(y_{M_Y-1}|x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_{Y|X}(y_0|x_{M_X-1}) & p_{Y|X}(y_1|x_{M_X-1}) & \cdots & p_{Y|X}(y_{M_Y-1}|x_{M_X-1}) \end{bmatrix}$$

Sometimes, a special kind of graph or *trellis* diagram, such as that in Figure 2, is used to specify the transition probabilities instead of the channel matrix. In such case, the transition probabilities are given as the labels of the edges joining the input and output elements. By definition, the probabilities of all the edges leaving a certain input node add up to 1.

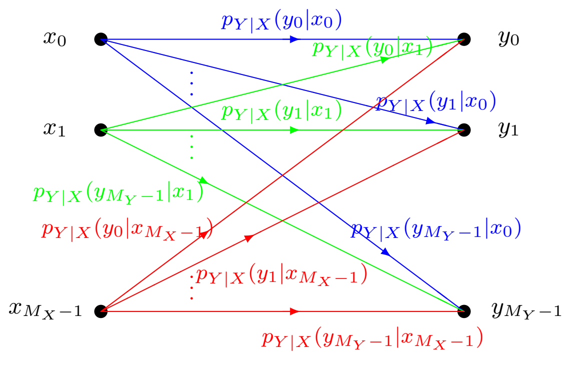

Figure 2: Representation of a discrete memoryless channel through a trellis diagram.

1.3 Capacity of a channel for discrete channels

Formally, the capacity, C, of a discrete memoryless channel is defined as the maximum value of the mutual information between the input and the output accounting for all the possible input distributions.

$$C = \max_{p_X(x_i)} I(X,Y).$$

It is measured in bits per channel use.

2 Estimation of quantiative metrics of information

This section revolves around estimating several quantitative information metrics from realizations of random variables. Specifically, we will be using the realizations of random variables X and Y within file

XYvariablesData.mat

After loading from the above file the values of the joint realizations of the random variables, you must perform the following tasks

1. Find the probability distributions of the random variables X and Y (probabilities $p_X(x_i)$ and $p_Y(y_j)$), and also the joint probability distribution $p_{XY}(x_i, y_j)$. Use the provided function

   ```
   [aX,aY,prX,prY,prXY]=estimateProbabilities(X,Y);
   ```

2. Implement a function named *estimateEntropy* that yields the estimate of the entropy of a given random variable. The prototype of the function must be

   ```
   [HX]=estimateEntropy(prX);
   ```

3. Implement a function named *estimateConditionalEntropy* that yields the estimate of the conditional entropy of X given Y, $H(X|Y)$ and that of Y given X, $H(Y|X)$. The prototype of the function must be

   ```
   [HXcY,HYcX]=estimateConditionalEntropy(prX,prY,prXY);
   ```

4. Implement a function named *estimateJointEntropy*. The prototype of the function must be

   ```
   [HXY]=estimateJointEntropy(prXY);
   ```

5. Implement a function named *estimateMutualInformation* that, from the distributions of X and Y, along with their joint distribution, computes the mutual information. The prototype of the function must be

   ```
   [I]=estimateMutualInformation(prX,prY,prXY);
   ```

6. Answer the questions sheet using the implemented functions.

3 Numerical estimation of the channel capacity

Use the provided function for computing the capacity

```
[Cmax,pXmax]=estimateCapacity(PYcX);
```

and answer the questions in the survey for the channel matrices, $\mathbf{P}(Y|X)$, there given.