Chapitre 1

Mesure de l'information

I.1 Introduction

La théorie de l'information est introduite comme moyen d'étudier et de résoudre les problèmes de communication ou de transmission de signaux à travers des canaux. En effet elle s'intéresse aux moyens de transmettre une information depuis une source jusqu'à un utilisateur.
L'étude des systèmes de communication est basée sur la séparation des modèles de sources et des modèles des canaux. Ceci peut se schématisé en séparent le codeur et le décodeur du système de communication. La figure 1 présente cette séparation qui permet de traiter le codage source et le codage canal indépendamment.

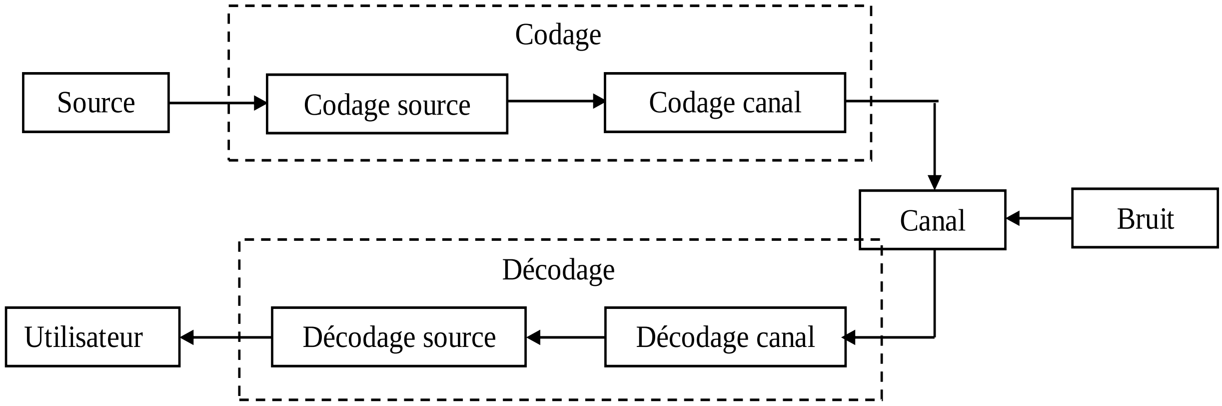


Figure 1 : Système de communication

La source est le siège d'événement aléatoire qui constitue le message émis. Sa nature est très variée, elle peut être une voix, un signal électromagnétique, une séquence de symboles….
Le canal permet de transmettre le message émis. Il est généralement perturbé par un bruit qui dépendra de l'environnement et de sa nature. Sa nature peut être une ligne téléphonique, une liaison radio ou encore un support magnétique ou optique.
Le codage source est d'attribuer une séquence binaire pour chaque symbole de l'alphabet. En effet, le code obtenu doit être uniquement décodable et efficace.
Le codage canal est d'ajouter des bits supplémentaires de contrôle qui permet de protéger le signal contre les perturbations.

I.2 Théorie des probabilités discrètes.

Une source d'information est décrite comme un objet qui produit un événement dont l'occurrence est déterminée d'une manière aléatoire. Dans ce contexte, La théorie d'information qui décrit les aspects les plus fondamentaux des systèmes de communication, nécessite pour son étude et son interprétation l'intégration de modèles mathématiques à l'aide essentiellement de la théorie des probabilités.

I.2.1 Probabilité conditionnelles.

La probabilité conditionnelle d'un événement X étant donné un événement Y, noté P(X/Y) est définit par $P(X/Y) = \frac{P(X \cap Y)}{P(Y)} \quad avec \; P(Y) > 0$.

La probabilité conditionnelle d'un événement Y étant donné un événement X, noté P(Y/X) est définit par $P(Y/X) = \frac{P(Y \cap X)}{P(X)} \quad avec \; P(X) > 0$

$$\Rightarrow P(X \cap Y) = P(Y \cap X) = P(X/Y)P(Y) = P(Y/X)P(X)$$

$$\Rightarrow \text{Première règle de bayes} : P(X/Y) = \frac{P(Y/X)P(X)}{P(Y)}$$

I.2.2 Probabilité totales

Soi X={X_1, X_2, ……,X_m} un système complet d'événements , alors soit Y un événement quelconque, on définit la probabilité totale de Y par :

$$P(Y)=\sum_{i=1}^{m}P(Y\cap X_i)=\sum_{i=1}^{m}P(Y/X_i)P(X_i)$$

$$\Rightarrow \textit{Deuxième}\ \text{règle de bayes}: P(X_i/Y)=\frac{P(Y/X_i)P(X_i)}{\sum_{i=1}^{m}P(Y/X_I)P(X_I)}$$

I.3 Mesure de l'information

I.3.1 Incertitude et information

La façon la plus appropriée de décrire un système de communication est de donner un modèle probabiliste. En effet, fournir une information à un utilisateur consiste à choisir un événement parmi plusieurs possibles, donc mesurer une incertitude sur l'issue d'une expérience aléatoire. De ce fait, pour mesurer la variation de l'incertitude et de décrire d'une manière quantitative il faudra choisir une fonction décroissante de la probabilité :

- Quand la probabilité de l'événement augmente ⇒ l'incertitude de l'événement diminue.
- Quand la probabilité de l'événement diminue ⇒ l'incertitude de l'événement augmente.

De même cette fonction doit satisfaire aux propriétés suivantes :

- $I(X) \geq 0$
- Si $P(X) = 1 \Rightarrow I(X) = 0$.
- Si X et Y sont deux événements indépendants, la quantité d'information doit être égale à la somme.

Le logarithme permet donc d'exprimer convenablement la variation d'incertitude dans un système de communication. $\Rightarrow I(X)=Log\frac{1}{P(X)}=-Log\ P(X)$

La base du logarithme détermine l'unité de mesure de l'incertitude en effet

- Si la base du logarithme est e, l'unité de mesure de I(X) est le Nat.
- Si la base du logarithme est 2, l'unité de mesure de I(X) est le Bit.
- Si la base du logarithme est 10, l'unité de mesure de I(X) est le Decit.

I.3.2 Information propre et information mutuelle

L'information propre présente la quantité d'information fournie par la réalisation d'un évènement :

$$I(X)=Log_2\frac{1}{P(X)}=-Log_2\ P(X)$$

L'information mutuelle présente une mesure quantitative de ce qu'apporte la réalisation d'un évènement Y sur la réalisation d'un évènement X. L'information mutuelle est définit par :

$$I(X,Y)=Log_2\frac{P(X/Y)}{P(X)}$$

$$\Rightarrow I(X,Y)=Log_2(P(X/Y))-Log_2P(X)=-Log_2P(X)-(-Log_2(P(X/Y)))$$

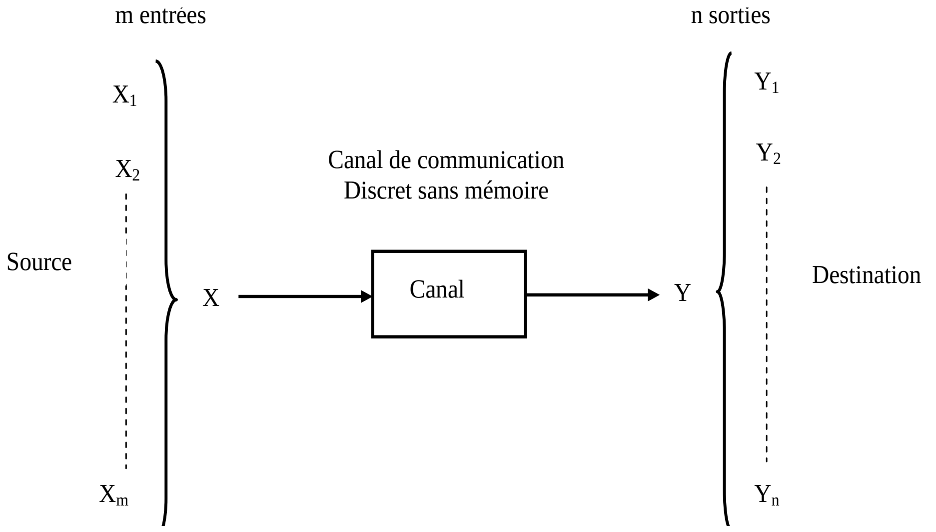
⇒ Information Mutuelle= information Propre- Information Conditionnelle

On a $P(X\cap Y)=P(X/Y)P(Y)=P(Y/X)P(X)\Rightarrow I(X,Y)=I(Y,X)=Log_2\frac{P(X\cap Y)}{P(X)P(Y)}$

- Si X et Y sont indépendants on a :
  $P(X\cap Y)=P(X)P(Y) \Rightarrow I(X, Y)= 0$.
- Si P(X/Y) = 1 on a :
  $I(X,Y)=-Log_2\ P(X)$: Information propre.

I.4 Information moyenne ou entropie

X est une source sans mémoire définie par son alphabet $X=\{X_1, X_2, \ldots\ldots\ldots, X_m\}$ et ses caractéristiques d'émission régies par une loi de probabilité P : $P(Xi)= \{P(X_1), P(X_2), \ldots\ldots.., P(X_m)\}$.
Y est une destination définie par son alphabet $Y=\{Y_1, Y_2, \ldots\ldots\ldots, Y_n\}$ et ses caractéristiques de réception régies par une loi de probabilité P : $P(Yj)= \{P(Y_1), P(Y_2), \ldots\ldots.., P(Y_n)\}$.



Avec

- $P(Xi)$: les probabilités d'entrées
- $P(Yj)$: les probabilités de sorties
- $P(Yj/Xi)$: les probabilités de transitions
- $P(Yj, Xi)$: les probabilités conjointes

Nous pouvons donc définir plusieurs fonctions d'entropies pour un canal de communication

I.4.1 Entropie propre.

L'entropie propre présente la quantité d'information moyenne associée à chaque symbole, elle est définie comme l'espérance mathématique de l'information propre fournie par l'observation de chacun des symboles possible. Elle est notée :

Source : $H(X) = -\sum_{i=1}^{m} P(X_i) Log_2(P(X_i))$

Destination : $H(Y) = -\sum_{j=1}^{n} P(Y_j) Log_2(P(Y_j))$

I.4.2 Entropie conditionnelle

H(X/Y) est l'entropie conditionnelle qui représente la mesure de l'incertitude moyenne des entrées étant donnée les sorties. Elle est notée par :

$$H(X/Y) = -\sum_{i=1}^{m}\sum_{j=1}^{n} P(X_i, Y_j) Log_2(P(X_i/Y_j))$$

H(Y/X) est l'entropie conditionnelle qui représente la mesure de l'incertitude moyenne des sorties étant donnée les entrées. Elle est notée par :

$$H(Y/X) = -\sum_{i=1}^{m}\sum_{j=1}^{n} P(Y_j, X_i) Log_2(P(Y_j/X_i))$$

I.4.3 Entropie mutuelle

H(X, Y) est l'entropie mutuelle qui représente la mesure de l'incertitude moyenne du canal de communication comme un tout. Elle est notée par :

$$H(X,Y) = -\sum_{i=1}^{m}\sum_{j=1}^{n} P(X_i, Y_j) Log_2(P(X_i, Y_j))$$

I.4.4 Relation entre les différentes entropies

La figure 2 présente une manière simple d'exprimer les différentes relations entre les distinctes entropies d'un système de communication :

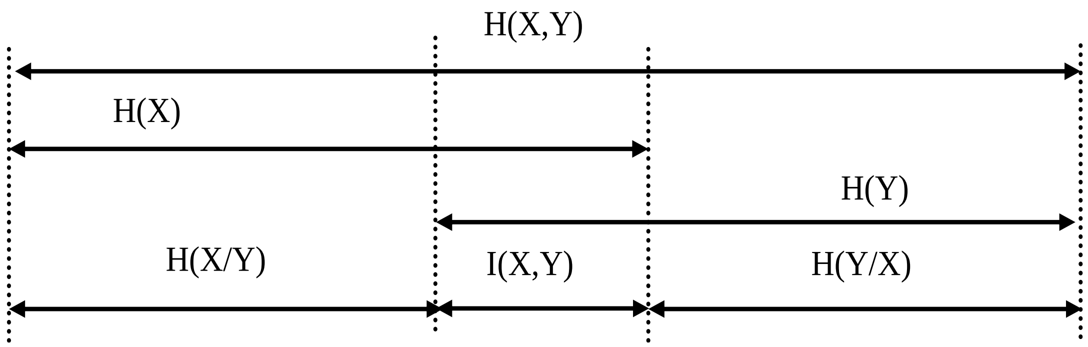
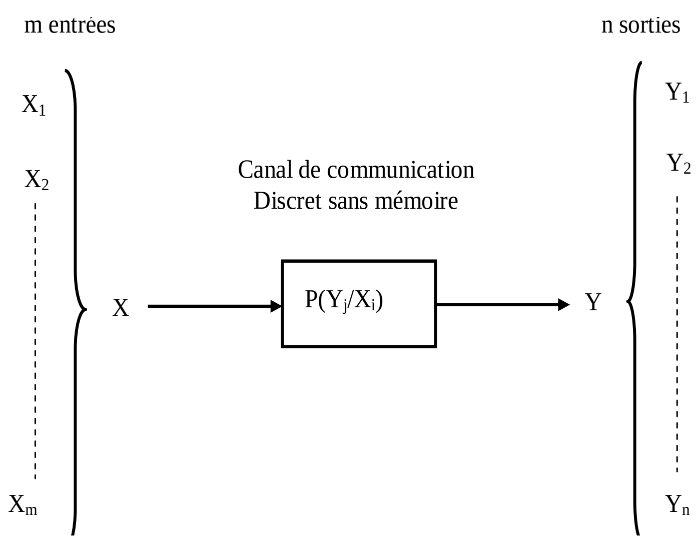


Figure 2 : différentes relations entre les distinctes entropies d'un système de communication

$$H(X,Y) = H(X/Y) + H(Y/X) + I(X,Y)$$
$$H(X) = H(X/Y) + I(X,Y)$$
$$H(Y) = H(Y/X) + I(X,Y)$$
$$H(X,Y) = H(X) + H(Y/X)$$
$$H(X,Y) = H(Y) + H(X/Y)$$

I.5 Canal de communication discret sans mémoires

Un canal de communication discret sans mémoire est la voie sur laquelle transitent les symboles vers le récepteur. Pour modéliser correctement un canal de transmission il est nécessaire de spécifier l'ensemble des entrées et l'ensemble des sorties possibles. La figure 3 présente un canal de communication discret sans mémoire qui est un modèle stochastique avec une entrée X et une sortie Y. Chaque lien entrée sortie est indiqué par une probabilité conditionnelle $P(Y_j/X_i)$, cette probabilité est appelée la probabilité de transition du canal de communication.



I.5.1 Matrice du canal de communication

L'étude des différentes entropies d'un système de communication nécessite la connaissance des différentes probabilités (propres, conditionnelles et conjointes). En effet, un canal de communication discret sans mémoire est caractérisé par l'ensemble complet des probabilités de transition qui représente ces caractéristiques techniques. Ces probabilités conditionnelles sont groupées dans une matrice dite matrice des de transition d'un canal de communication. La connaissance de cette matrice permet de déterminer les probabilités de sorties sachant les probabilités d'occurrence des symboles de la source. Elle permet aussi de déterminer les différentes probabilités conjointe du système. Cette matrice est notée par :

$$[P(Y/X)] = \begin{bmatrix} P(Y_1/X_1) & P(Y_2/X_1) & . & . & P(Y_n/X_1) \\ P(Y_1/X_2) & P(Y_2/X_2) & . & . & P(Y_n/X_2) \\ . & . & . & . & . \\ . & . & . & . & . \\ P(Y_1/X_m) & P(Y_2/X_m) & . & . & P(Y_n/X_m) \end{bmatrix}$$

Chaque ligne de la matrice du canal de communication doit avoir une somme égale à 1. En effet on a :

$$\sum_{j=1}^{n} P(Y_j/X_i) = 1$$

Les probabilités des entrées et des sorties sont présentées aussi par des matrices :

$$[P(X)] = [P(X_1\ \mathrm{P(X_2)}\ldots\ldots\ldots\ldots\mathrm{P(X_m)}]$$

$$[P(Y)] = [P(Y_1\ \mathrm{P(Y_2)}\ldots\ldots\ldots\ldots\mathrm{P(Y_n)}]$$

La détermination des probabilités des sorties sachant la matrice de transition du canal de communication est assurée par l'équation suivante :

$$[P(Y)] = [P(X)] * [P(Y/X)]$$

La détermination des différentes probabilités conjointes du système de communication est assurée par la relation suivante :

$$[P(X,Y)] = [P(X)]_d * [P(Y/X)]$$

Avec $[P(X)]_d = \begin{bmatrix} P(X_1) & 0 & . & . & 0 \\ 0 & P(X_2) & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & . & . & . & P(X_m) \end{bmatrix}$ matrice diagonale

I.5.2 Capacité d'un canal

Un canal de communication ou canal de transmission est un support physique permettant la transmission d'une certaine quantité d'information, depuis une source vers un destinataire. La capacité d'un canal de communication représente le taux de transmission maximal du code à utiliser pour transmettre de l'information dans de bonne condition. Dans ce contexte la capacité d'un canal de communication discret sans mémoire est définie par :

$$C = \underset{P(X_i)}{Sup}\ \mathrm{I(X,Y)} \quad \text{avec } \mathrm{I(X,Y)} = H(X) - H(X/Y)$$