Chapitre 2 : Mesure de l'information

Avant de trouver un moyen de mesurer l'information, nous allons essayer de préciser ce concvept :

1. **Définition de l'information** : c'est un concept ayant plusieurs sens :
   - Désigne à la fois le message à communiquer et les symboles utilisés pour l'écrire ;
   - Séquence de signaux, transmise entre une source et un destinataire par l'intermédiaire d'un canal ;
   - Une grandeur observable et mesurable, elle présente un caractère aléatoire.

2. **Incertitude et Information :**

   Emmetre un message revient à émettre une succession de symboles appartenet à une source.

   Exemple :

   Bonjour → une succession de symboles B, o, n, j, o, u, r, ces symboles appartient à l'alphabet français.

   Donc, une source dispose d'un alphabet constitué d'éléments ou symboles ou caractères {x1, x2, …, xn} :

   Où : n est la longueur de l'alphabet ;

   Ces symboles sont associés pour constituer un message ;

   Chaque symbole de l'alphabet possède one probabilité d'utilisation p.

   ➔ Si on reçoit un message qu'on a déjà connu à l'avance son contenu : il ne nous apportye aucune information ou nouvelle.

   ➔ Un message reçu apporte des nouvelles si son contenu n'est pas connu à l'avance.

   ➔ Une information est ainsi plus riche qu'elle est peu probable.

   P ↗ → quantité d'information ↘

   P ↘ → quantité d'information ↗

   ➔ La quantité d'information est noté I .

2.1 La quantité d'information

- Soit une source S qui envoi des symboles {x1, x2, …, xn} de manière aléatoire, P(xi) est la probabilité d'envoi du symbole xi ;

On a $\sum_{i=1}^{n} p(xi)$=1

- Lorsque l'évènement S=xi se réalise il produit une information, cette information est liée à la probabilité de cet évènement.
- La quantité d'information obtenu lorsque l'évènement S=xi se réalise est liée à l'incertitude sur cet évènement.

  Donc nous chercherons :
- Une fonction positive, et décroissante de la probabilité : I(x)=f(p(x))

  P(x) ↗ I(x) ↘ et vise versa
- L'évènement certain ne produit aucune information : I(x)=f(p(x)=1)=0.
- L'évènement impossible fournit une quantité infinie de l'information : I(x)=f(p(x)=o)= ∞
- Une fonction additive, l'information de deux évènements indépendants s'additionne : I(x, y)=f(p(x) p(y))= f(p(x))+ f(p(y)).

➔ Le **logarithme** permet d'exprimer l'incertitude, donc pour mesure l'incertitude, nous utiliserons la quantité suivante :

I(x)=$\log_b(\frac{1}{p(x)})$ l'unité est le bit.

b : représente le nombre de symboles nécéssaires pour le codage de source (généralement nous utilisons le bit (0,1))➔ le logarithme en base 2).

$$\boxed{I(x) = -\log_2(p(x))}$$

I(x) sera appelée « **information propre** » de x, et s'interprète aussi comme la quantité d'information fournie par la réalisation d'un événement.

2.2 Entropie d'une source

L'entropie H(S) d'une source, est la quantité d'information moyenne contenu dans l'alphabet S de cette source.

H(S)=$\sum_{i=1}^{n} p(xi).\log_2 \frac{1}{p(xi)}$

H(S)= $\boldsymbol{-\sum_{i=1}^{n} p(xi).\log_2 p(xi)}$ l'unité est bits/symboles

Cette entropie admet une limite maximale H(S)$\leq \log_2 n$, elle est atteinte lorsqu'il ya une équiprobabilité entre tous les symboles de l'alphabet de la source.

2.3 Entropie conditionnelle :

L'entropie conditionnelle de S sachant T est définie par :

$$\mathbf{H(S/T)} = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(xi, yj) \log_2 p(^{xi}/_{yj})$$

$-\log_2 p(^{xi}/_{yj})$: est l'information propre conditionnelle de xi sachant yj : I($^{xi}/_{yj}$).

xi ϵ **S=**{x1, x2, …, xn}

yj ϵ **T=**{y1, y2, …, ym}

3. L'information mutuelle :

Soient x,y deux évènements, nous désirons donner une mesure quantitative de ce que apporte la réalisation d'un évènemnt sur la probabilité de réalisation de l'autre évènement (c'est-à-dire quantifier la correlation entre les deux évènements) ?

I(x ; y)= ?

- L'information mutuelle entre x et y est égale à la quantité d'information fournie par x moins la quantité d'information que fournirait x lorsque y est réalisé.

I(x ; y)= I(x) – I(x/y)

$$= -\log_2 p(x) - (-\log_2 p(x/y))$$

$$= \log_2 p(x/y) - \log_2 p(x) = \log_2 \frac{p(x/y)}{p(x)} = \log_2 \frac{p(x,y)}{p(x)p(y)} = \log_2 \frac{p(y,x)}{p(x)p(y)} = I(y ; x)$$

une quantité symétrique

$$\boxed{\mathbf{I(x ; y)} = \log_2 \frac{p(x,y)}{p(x)\, p(y)}}$$

- I(x ; y) > 0 SSi p(x/y) > p(x) : la réalisation de y augmente la probabilité d'occurrence de x.
- I(x ; y) < 0 SSi p(x/y) < p(x) : la réalisation de y diminue la probabilité d'occurrence de x.
- I(x ; y) = 0 SSi p(x/y) = p(x) : les deux évènements sont indépendants.

4. L'information mutuelle moyenne :

Soient S et T deux sources discrètes où :

S={x1, x2, …, xn}

T={y1, y2, …, ym}

L'information mutuelle moyenne de S et T dans la source ST est définie par :

I(S ; T)=$\sum_{i=1}^{n}\sum_{j=1}^{m} p(xi, yj)\ I(xi; yj)$

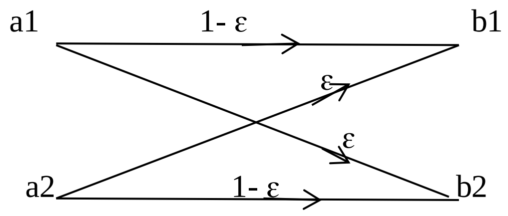
I(S ; T)=$\boldsymbol{\sum_{i=1}^{n}\sum_{j=1}^{m} p(xi, yj) \log_2 \frac{p(xi,yj)}{p(xi)p(yj)}}$

On a aussi :

I(S ; T)= H(S)-H(S/T)=H(T)-H(T/S)=I(T ;S)

Cette relation nous sera utile pour interpréter la capacité du canal.

Exemple : considérons le canal binaire symétrique de probabilité ε avec les entrés a1 et a2 équiprobables.



Le canal est défini par les probabilités conditionnelles :

p(b1/a1) = P(b2/a2) =1- ε

p(b1/a2) = P(b2/a1) = ε

Puisque p(a1)=p(a2) =1/2, on en déduit :

p(a1, b1) = p(a2, b2) = $\frac{1-\varepsilon}{2}$

p(a1, b2) = p(a2, b1) = $\frac{\varepsilon}{2}$

p(b1)=p(b2)=1/2

Cela nous permet de calculer l'information mutuelle de chaque couple (a_k,b_k)

I(a1 ; b1) = I(a2 ; b2) = $\log_2 2(1-\varepsilon)$

I(a1 ; b2) = I(a2 ; b1) = $\log_2 2\varepsilon$

- Si $\varepsilon < \frac{1}{2}$ → I(a1 ; b1)>0 et I(a1 ; b2)<0

  Si on observe à la sortie du canal la lettre b1, la probabilité pour que a1 ait été émisse augmente.

- Si $\varepsilon = \frac{1}{2} \rightarrow$ toutes les informations mutuelles sont nulles, et donc l'alphabet d'entrée et de sortie sont statistiquement indépendants, ce qui n'est évidement souhaitable.

5. Entropie et Information dans le cas continu

5.1 Information propre

$$I(x) = -\log_2(p(x))$$

5.2 Information mutuelle

$$I(x\,;\,y) = \log_2 \frac{p(x,y)}{p(x)\,p(y)}$$

5.3 Information propre conditionnelle de x sachant y

$$I(x/y) = -\log_2(p(x/y))$$

5.4 Information mutuelle moyenne

$$\mathbf{I(S\,;\,T)} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \boldsymbol{p(x,y)} \,\mathbf{log_2} \frac{\boldsymbol{p(x,y)}}{\boldsymbol{p(x)p(y)}} \boldsymbol{d(x)d(y)}$$

5.5 L'entropie

$$\mathbf{H(S)} = -\int_{-\infty}^{+\infty} \boldsymbol{p(x)} \,\mathbf{log_2}\, \boldsymbol{p(x)d(x)}$$

5.6 L'entropie conditionnelle de S sachant T

$$\mathbf{H(S/T)} = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \boldsymbol{p(x,y)} \,\mathbf{log_2}\, \boldsymbol{p(x/y)d(x)d(y)}$$