Théorie de l'Information et Codage

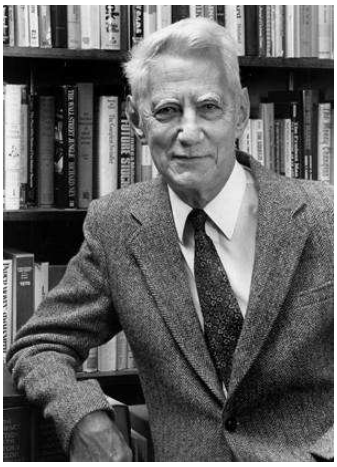
M. Belkasmi

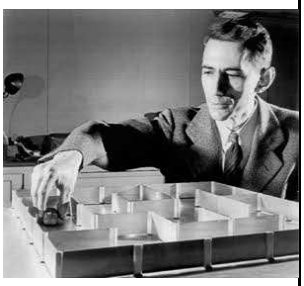
2012-2013 ENSIAS

Plan du cours

- ✓ Introduction à la théorie de l'information
- ✓ Mesure de l'information et entropie.
- ✓ Codage de source.
- ✓ Cryptographie.
- ✓ Canaux de transmission et codage.
- ✓ Codage de canal
- ✓ Codes en bloc
- ✓ Codes Convolutionnels
- ✓ Turbo codes et Décodage itératif

Shannon (1948) , Théorie de l'Information,
The Mathematical theory of Communication





Claude Elwood Shannon: April 30, 1916 - February 24, 2001

Bibliographie :

-J.Clavier et al : *Théorie et technique de la transmission des données* , tome 1, Masson, 1989.

-A. Poli et al.: *Codes correcteurs théorie et applications*, Masson, 1989

-H. NUSSBAUMER « Téléinformatique », Presses polytechniques Romandes, 1987.

-G. COHEN et al. « Codes Correcteurs d'erreurs », Masson, 1992.

- A. TANENBAUM « Réseaux, Architecture,… » InterEdition 1990.

Chapitre 1 : Introduction à la Théorie de l'Information

INFORMATION : Communication ou réception de renseignements, données, faits nouveaux, connaissances résultant de l'étude d'une observation.

Comment bâtir une théorie scientifique de l'information ?

Un pb : → un certain nb de solutions possibles (lorsqu'on ne possède pas d'info. sur la situation présente)

- **Si info suppl. sur le pb → nb de réponses possibles ↘**
- **Si info totale → une seule réponse possible**

- → L'info : Rapport du nb de réponses possibles après et avant Réception.
- ↗ Si elle lève au mieux l'ambiguïté

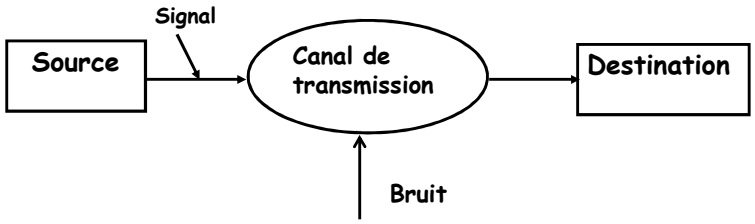
Domaines D'application :

- Codage, télécom, parole, ...
- Ici, on transforme et/ ou transmet l'information d'un point à un autre.
- La T.I. :
  - * précise ce qui est possible et impossible.
  - * Scientifique et mesurable.

Historique :

- 1928: 1ere initiative de définition de quantité d'information (HARTLEY)
- 1948: naissance de la TI avec les travaux de SHANNON
- 1950s : Construction de plusieurs familles de codes
- 1960s: Algorithmes pour correction d'erreurs et compression de données

Représentation d'un système de communication



. Canal ou voie = transporteur d'information.

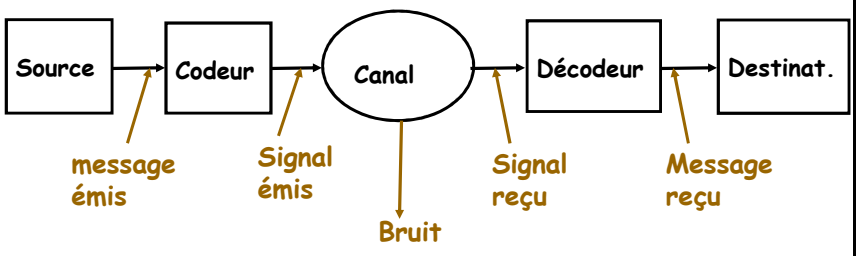
- Bruit : → l'information chez l'utilisateur incomplète, distordue.
- Action : adapter le signal

Action : adapter le signal

→ une forme plus assimilable par la voie = « codage »

à l'arrivée : traitement réciproque = « décodage »

But : rendre le S.C. le plus efficace possible.



Relation avec les probabilités

Info d'une source : ensemble de N messages.

- Les N messages sont 'possibles'
  → chacun a une probabilité.
  (Si aucune incertitude sur le message émis → pas d'info à la réception)

- Apport d'information ≡ Réduction d'incertitude.

Exemple

Source :

- délivre une lettre / T sec de façon indépendante
- parmi { a, b, c, d, e, f, g, h }
- apparition :

$p_a = 1/4$ $p_b = 1/4$ $p_c = 1/8$ $p_d = 1/8$
$p_e = 1/16$ $p_f = 1/16$ $p_g = 1/16$ $p_h = 1/16$

- proba(h) < proba(a) → récepteur apprend plus avec h qu'avec a

⇒ Info plus grande ↔ événement de plus faible proba.

Exemple suite

Source seconde vision :

- Toutes 2T sec → $8^2 = 64$ événements possibles (ab, bc, ….)

- Si I1 = info, 1ere période T, } et I12 = info dans 2T
  I2 = info, 2eme période T, }

⇒ I1 + I2 = I12 : additivité de l'info

Synthèse

- Apport d'information ≡ Réduction d'incertitude.
- Info plus grande ↔ événement de plus faible proba.
- Additivité de l'info

Chapitre 2 :
MESURE DE L'INFORMATION ET ENTROPIE

<u>Définition de l'information</u> :

- Une expérience [X] (champs d'événements) ayant N résultats possibles :

  [X] = [x1, x2, ..., xN] ,

- chaque résultat xi → proba p(xi)

  [P] =[p(x1), p(x2), …, p(xN)]

- mesure incertitude réalisation xi : fonction proba. a priori p(xi)

Définition :

L'information résultat xi (self-info fournie par xi):

I(xi) = F (p(xi)) : mesure de l'incertitude a priori

On montre que :

. F additive F (p(z1)p(z2)) = F(p(z1)) + F(p(z2))

. F continue, ≥ 0 , monotone

→ F ne peut être que le Logarithme

→ I(xi) = -λ Log p(xi); λ facteur d'échelle > 0

Unités de mesure de l'information

L'expérience la plus simple possible (pile ou face)

Définition :

1 bit (ou 1 shannon) = quantité d'information correspondant à l'un des 2 résultats .

I(x1)= I(x2) = - λ log $\frac{1}{2}$ = 1 bit 'BInary uniT'

si base du log = 2 → λ =1

. Si Log_e = Ln → I exprimée en nat
'NAtural uniT'

. Si Log_{10} → I exprimée en decit
'DECimal uniT'

Pour la suite de ce cours nous choisissons unité = bit.

Exemples

- Exemple précèdent :
  I(a)=I(b)= 2 bits,
  I(c)=I(d)= 3 bits,
  I(e)= I(f)= I(g)=I(h)= 4 bits.

- Quantité d'information associée à un résultat certain :
  p =1 → I = 0 ce qui est logique

Exemples (suite)

• Quantité d'information associée à un résultat pris parmi N équiprobables :

[X] = [x1, x2, ... xN] , proba p(xi) = 1/N

I(xi)= - log 1/N = log N bits quelque soit i

En particulier, si N = 2^K alors I(xi)= K bits.

Modèles de sources discrètes

Une Source est définie par deux quantités :

- un ensemble fini de messages ou symboles :
[X] = [x1, ..., xN] (appelé encore 'alphabet')

- un mécanisme d'émission de suites de tels messages suivant une loi de probabilité donnée.

Modèles de sources discrètes

Il est commode de considérer que les messages successifs d'une suite :

$$S_n = x_{\alpha 1}\, x_{\alpha 2} \ldots x_{\alpha n}$$

ont été émis à des instants notés 1, 2, … , n.

La situation est celle des processus discrets :

Une source est l'équivalent d'une suite de variables aléatoires (v.a.) X1, X2, …, Xn à valeurs dans [X] et ayant des lois de probabilité données.

Les Différentes sources

1. source sans mémoire (simple)

Les v.a. Xk sont indépendantes et de même loi

$$P(X_k = x_i) = p(x_i) =^{def} p_i$$

pour tout k. Ainsi

$$\mathrm{Prob}(S_n) = p_{\alpha_1}\, p_{\alpha_2} \cdots p_{\alpha_n}$$

Les Différentes sources

2. source stationnaire

La loi conjointe de k v.a. est invariante par translation dans le temps :

$$P[X_1 = x_{\alpha_1}, \cdots, X_k = x_{\alpha_k}] = P[X_{1+h} = x_{\alpha_1}, \cdots, X_{k+h} = x_{\alpha_k}]$$

pour tout k, $\alpha 1$, …, αk et h

Les Différentes sources

3. source de Markov

La mémoire du passé est résumée dans les r derniers messages où r est un nombre entier assez petit.

4. sources quelconques

Les v.a. Xk sont conditionnées les unes par les autres. C'est un cas assez compliqué.

Entropie d'une source simple

Définition

Une source simple S caractérisée par [x1,.., xN], [p1,.., pN].
L'entropie de cette source est l'espérance mathématique de la v.a. numérique I(X) dont les réalisations possibles sont les I(xi)=-logp(xi)

$$H(X) = -\sum_{i=1}^{N} p(x_i) \log p(x_i)$$

Ceci correspond à une information moyenne par message de la source.

Exemple 1

Source à N messages équiprobables : pi = 1/N quelque soit i

→ I(xi) = log N bits

$$H(X) = -\sum_{i=1}^{N} \frac{1}{N} \log \frac{1}{N} = \log N \text{ bits}$$

Symétrie de la loi → l'information associée à chaque message est la même que l'information moyenne.

→ Pour une expérience de pile ou face, H = 1 bit.

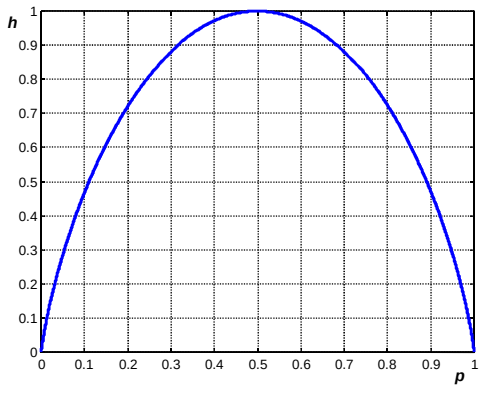
Exemple 2

Une source binaire peut délivrer deux messages possibles, par exemple 0 et 1, avec des probabilités p et 1-p.

L'entropie d'une telle source :

$$H(p) = -p\log(p) - (1-p)\log(1-p)$$

Courbe de l'entropie binaire



→ H(p) s'annule pour p = 0 et 1 et est maximale pour p = 1/2, auquel cas elle prend pour valeur 1 bit /lettre.

→ le bit est la quantité d'information qui correspond au lever du doute entre 2 messages équiprobables.

Propriétés de l'entropie

. L'entropie H(p1,...,pN) est une fonction de N variables. Ces variables sont liées par la relation :

$$\sum_{i=1}^{N} p_i = 1$$

- Cette fonction est continue par rapport à chacune des variables pk dans l'intervalle [0,1].
- Symétrique par rapport à toutes les variables p1, p2,.....,pN.
- Entropie bornée : $0 \leq H(p1, ..,pN) \leq \log N$.