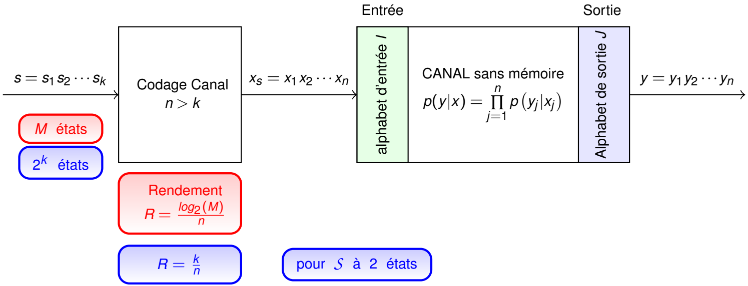
Second Théorème de Shannon

Notations

- X_n l'ensemble des suites x de longueur n en entrée de canal
- Y_n l'ensemble des suites y de longueur n en sortie de canal

Hypothèses

- $\mathcal{S}$ souce simple et de loi uniforme
- Le canal est sans mémoire à entrées et sorties discrètes



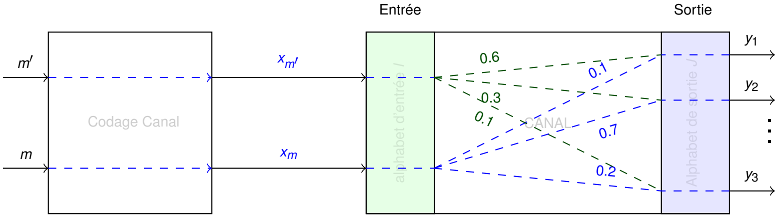
Dans le cas d'une source simple à deux états et d'un codage binaire : 2^k mots de codes choisis parmi 2^n

Décodage au sens du maximum de vraisemblance (voir TD 8)

Connaissant la séquence y en sortie le mot x_m en entrée est vraisemblablement celui pour lequel $p(y|x_m)$ est maximal.
On décode donc y en choisissant le mot $x_m \in X_n$ tel que

$$(\forall x_{m'} \in X_n)\,(x_{m'} \neq x_m \Longrightarrow p(y|x_m) > p(y|x_{m'}))$$

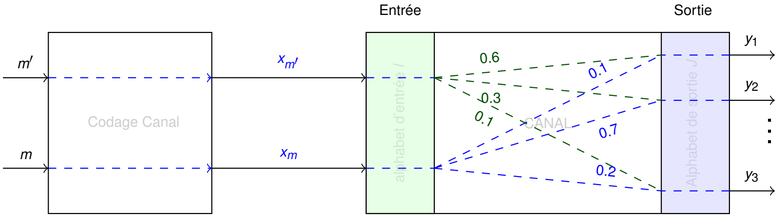
Probabilité d'erreur de décodage pour un mot donné en entrée



Exemple 1 : 2 mots en entrée, 3 mots possibles en sortie

- Sachant que x_m a été émis en entrée de canal, quelle est la probabilité d'avoir un décodage érroné ?
- Même question sachant que $x_{m'}$ a été émis en entrée de canal

Probabilité d'erreur de décodage pour un mot donné en entrée



Exemple 2 : en entrée tous les mots de X_n sont possibles, 3 mots possibles en sortie lorsque x_m est en entrée

- Sachant que m a été émis en entrée de canal, quelle est la probabilité d'avoir un décodage érroné ? On pourra utilisé la fonction Φ défini ci dessous :

$$\Phi_m(y) = \begin{cases} 1 & si\ il\ existe\ m'\ tel\ que\ p(y|x_{m'}) > p(y|x_m) \\ 0 & sinon \end{cases}$$

Majoration de la probabilité d'erreur de décodage pour un mot donné en entrée

$$\boxed{P_e(m) = \sum_{y \in Y_n} P(y|x_m)\Phi_m(y)}$$

pour tout $m \in X_n$, $s > 0$, $y \in Y_n$ on peut majorer $\Phi_m(y)$ par

$$\Phi_m(y) \leq \left(\frac{\sum_{m' \neq m} P(y|x'_m)^{\frac{1}{1+s}}}{P(y|x_m)^{\frac{1}{1+s}}} \right)^s$$

$$\boxed{P_e(m) \leq \sum_{y \in Y_n} \left\{ P(y|x_m)^{\frac{1}{1+s}} \left(\sum_{m' \neq m} P(y|x'_m)^{\frac{1}{1+s}} \right)^s \right\}}$$

Codage aléatoire

- les mots de codes sont issus de tirages indépendants selont une loi $P(x)$ définie sur l'ensemble des séquences d'entrées
- lorsque les mots de codes sont aléatoires, les $P(y|x_m)$ qui en dépendent sont des variables aléatoires indépendantes
- $P_e(m)$ est une une variable aléatoire dont nous allons majorer l'espérance et en déduire qu'il existe nécessairement un code pour lequel la probabilité d'erreur est aussi inférieure à la borne obtenue

Majoration de la probabilité d'erreur moyenne pour un ensemble probabilisé de codes

- pour tout mot m on a $E\left(P(y|x_m)^{\frac{1}{1+s}}\right) = \sum_{x \in X_n} P(x)(P(y|x)^{\frac{1}{1+s}}$
- pour tout $0 < s < 1$ on a

$$
\begin{aligned}
E(P_e(m)) &\leq E\left\{\sum_{y \in Y_n}\left[P(y|x_m)^{\frac{1}{1+s}}\left(\sum_{m' \neq m} P(y|x'_m)^{\frac{1}{1+s}}\right)^s\right]\right\} \\
&\underbrace{\leq}_{linéarité} \sum_{y \in Y_n} E\left\{\left[P(y|x_m)^{\frac{1}{1+s}}\left(\sum_{m' \neq m} P(y|x'_m)^{\frac{1}{1+s}}\right)^s\right]\right\} \\
&\underbrace{\leq}_{indépendance} \sum_{y \in Y_n}\left\{E\left[P(y|x_m)^{\frac{1}{1+s}}\right] E\left[\left(\sum_{m' \neq m} P(y|x'_m)^{\frac{1}{1+s}}\right)^s\right]\right\} \\
&\underbrace{\leq}_{x^s\, concave} \sum_{y \in Y_n}\left\{E\left[P(y|x_m)^{\frac{1}{1+s}}\right]\left[E\left(\sum_{m' \neq m} P(y|x'_m)^{\frac{1}{1+s}}\right)\right]^s\right\} \\
&\underbrace{\leq}_{linéarité} \sum_{y \in Y_n}\left\{E\left[P(y|x_m)^{\frac{1}{1+s}}\right]\left[\sum_{m' \neq m} E\left(P(y|x'_m)^{\frac{1}{1+s}}\right)\right]^s\right\}
\end{aligned}
$$

$$
\boxed{E(P_e(m)) \leq (M-1)^s \sum_{y \in Y_n}\left(\sum_{x \in X_n} P(x)P(y|x)^{\frac{1}{1+s}}\right)^{1+s}, \quad 0 < s < 1}
$$

Borne de la probabilité d'erreur moyenne pour un canal sans mémoire et une source simple

- la source est source simple : $P(x = x_1 x_2 \cdots x_n) = \prod_{k=1}^{n} P(x_k)$
- le canal est sans mémoire :$P(y|x) = \prod_{k=1}^{n} P(y_k|x_k)$

on en déduit pour $0 < s < 1$

$$E(P_e(m)) \leq (M-1)^s \sum_{y \in Y_n} \left[\sum_{x \in X_n} \prod_{k=1}^{n} p(x_k) P(y_k|x_k)^{\frac{1}{1+s}} \right]^{\frac{1}{1+s}}$$

En se ramenant aux alphabets I d'entrée et J de sortie du canal on a

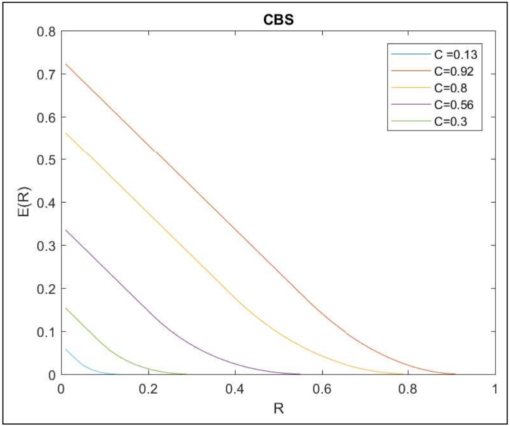
$$E(P_e(m)) \leq (M-1)^s \sum_{j \in J} \left[\sum_{i \in I} p(i) p(j|i)^{\frac{1}{1+s}} \right]^{1+s}, \quad 0 < s < 1$$

$$\boxed{\begin{array}{c} E(P_e(m)) \leq 2^{-nE(R)} \\ \text{où} \quad E(R) = max_{s,\{p(k)\}} \left\{ -sR - log_2 \left(\sum_{j \in J} \left[\sum_{i \in I} p(i) p(j|i)^{\frac{1}{1+s}} \right]^{1+s} \right) \right\} \end{array}}$$

Calcul de la borne pour un canal binaire symétrique pour une loi d'entrée uniforme

$$E\left(P_e(m)\right) \leq 2^{-nE(R)}$$

$$\text{où} \quad E(R) = max_{s\in]0,1[}\left\{s(1-R)-(1+s)log_2\left(p^{\frac{1}{1+s}}+(1-p)^{\frac{1}{1+s}}\right)\right\}$$



Commentaire sur le second théorème de Shanon

Si $R < C$ alors **il existe** un code canal de rendement R dont la probabilité d'erreur est aussi faible que souhaité sitôt à condition que la longueur des mots de codes soit suffisante
Il est possible de réduire la probabilité d'erreur à une valeur arbitrairement faible à rendement constant.

Structure $(\mathcal{B} = \{0,1\}, \oplus, \cdot)$ et $\mathcal{B}^k$

- $(\mathcal{B}, \oplus, \cdot)$ est un corps commutatif

$\oplus$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

- $(\mathcal{B}^k, \oplus, \cdot)$ est un $\mathcal{B}$-e.v.
  - $a_1 a_2 \cdots a_k \oplus b_1 b_2 \cdots b_k = (a_1 \oplus b_1)(a_2 \oplus b_2) \cdots (a_k \oplus b_k)$

    exemple : $\begin{array}{rcccccc} & 0 & 1 & 0 & 1 & 1 & 1 \\ \oplus & 1 & 1 & 0 & 0 & 0 & 1 \\ \hline & 1 & 0 & 0 & 1 & 1 & 0 \end{array}$
  - $\lambda.(b_1 b_2 \cdots b_k) = (\lambda.b_1)(\lambda.b_2) \cdots (\lambda.b_k)$

    exemples : 1. 1101101 = 1101101 et 0.1101101 = 0000000
  - vecteur (mot) nul $00 \cdots 0$
  - base canonique $\{e_i = a_1 a_2 \cdots a_k | a_i = 1 \; et \; j \neq i \Rightarrow a_j = 0, i = 1, \cdots, k\}$
  - remarque tout vecteur $= a_1 a_2 \cdots a_k$ est son propre opposé puisque $a \oplus a = 00 \cdots 0$
- "poids d'un mot " :

$$\omega(a_1 a_2 \cdots a_k) = \sum_{a_i = 1} a_i$$

La base canonique de $(\mathcal{B}^k, \oplus, \cdot)$ est l'ensemble des mots de poids 1

- distance de Hamming entre les mots $a = a_1 a_2 \cdots a_k$ et $b = b_1 b_2 \cdots b_k$

$$D_H(a,b) = \omega(a \oplus b)$$

$(\mathcal{B}^k, \prec)$ algèbre de Boole

- relation $\prec$ dans $\mathcal{B}$

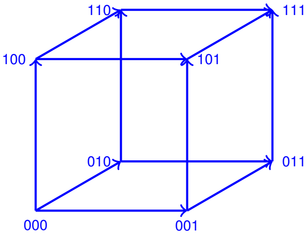
$$a \prec b \Longleftrightarrow a \leq b$$

- relation d'ordre $\prec$ dans $\mathcal{B}^k$

$$a_1 a_2 \cdots a_k \prec b_1 b_2 \cdots b_k \Longleftrightarrow (\forall j \in \{1, 2, \cdots, k\}) a_j \prec b_j$$

- $(\mathcal{B}^k, \prec)$ est une algèbre de Boole dont les atomes (successeurs immédiats du plus petit élément O) sont les mots de poids 1

Diagramme de Hasse de $(B^3, \prec)$



Dans cette représentation la distance de Hamming entre deux mots est le nombre minimale d'arêtes qui les séparent

Codage par bloc

un code par bloc de rendement $\frac{k}{n}$ est une application

$$\mathcal{C} : \begin{array}{ccc} \mathcal{B}^k & \to & \mathcal{B}^n \\ m & \to & C_m \end{array}$$

Notons C l'ensemble des mots du code $\mathcal{C}$

la distance du code $\mathcal{C}$ est la distance de Hamming minimale entre deux mots de codes distincts

$$d = min\{d_H(C_i, C_j), i \neq j, C_i \;\; et \;\; C_j \in C\}$$

Détection : on peut détecter $d-1$ erreur

Correction : on peut corriger de façon exacte selon la méthode du maximum de vraisemblance au plu