① Incertitude, Information, Entropie

- Incertitude de $X = x$: $\text{Incert}(x) = -\log_2(p(x))$ en bit
- Incertitude moyenne : Entropie de Shanon

$$H(X) = -\sum_{x \in \mathcal{A}} p(x) \log_2(p(x))$$

- Questionnaire sur X
  - ↳ # de questions $N \geq \lceil H(X) \rceil$
  - ↳ Questionnaire binaire optimal : Découpage en ensemble $\simeq$ équiprobable.

↳ maximisant / minimisant : $0 \leq H(p_1, \dots, p_N) \leq \log_2(N)$ bits

↳ $H = 0 \Leftrightarrow X$ déterministe

↳ $H = \log_2 N \Leftrightarrow X$ est uniforme : $(p_i) = \frac{1}{N}$

② Transinformation

- Entropie conjointe $H(X,Y) = -\sum_{x \in \mathcal{A}_X} \sum_{y \in \mathcal{A}_Y} p((x,y)) \log_2\left(p((x,y))\right)$ bit

  ↳ $0 \leq H(X,Y) \leq H(X) + H(Y)$

- Entropie Conditionnelle $H(X|Y) = \sum_{x \in \mathcal{A}_X} \sum_{y \in \mathcal{A}_Y} p(y) H(X|Y=y)$

avec $H(X|Y=y) = -\sum_{x \in \mathcal{A}_X} p(x|y) \log_2\left(p(x|y)\right)$

or $p(x|y) = \frac{p(x,y)}{p(y)}$ d'où $H(X|Y) = -\sum_{x \in \mathcal{A}_X} \sum_{y \in \mathcal{A}_Y} p(x,y) \log_2\left(p(x|y)\right)$

↳ $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$

↳ $H(X|Y) \leq H(X)$

- **Divergence de Kullback** (entropie relative)

$$D(p\|q) = \sum_{x\in A} p(x)\log_2\left(\frac{p(x)}{q(x)}\right) \geq 0$$

- **Information mutuelle**

$$I(X;Y) = D\big(p(x,y)\,\|\,p(x)p(y)\big) = \sum_{x\in A_X}\sum_{y\in A_Y} p(x,y)\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

↳ $I(X;Y) = H(X) - H(X|Y) = I(Y;X)$

③ Compressibilité & Entropie

On veut coder la source avec $C : \underbrace{\mathcal{A}_X}_{\{x_1,\dots,x_N\}} \to \underbrace{\mathcal{D}^*}_{\{a_1,\dots,a_d\}}$ où $|\mathcal{D}^*| < |\mathcal{A}_X|$

$x \mapsto C(x)$

- **Compacité** $\ell = \sum_{x\in\mathcal{A}_X} p(x)\,\ell(x)$ où ℓ = longueur du mot code ↳ $\ell(C(x))$
- **Code non singulier** $\Leftrightarrow$ C est injective.
- **Code instantané** $\Leftrightarrow$ Décodage pas à pas possible

Inégalité de Kraft Un code instantané existe $\Leftrightarrow \sum_{i=1}^{N} d^{-l_i} \leq 1$

- **Efficacité d'un code** $0 \leq \text{Eff} = \dfrac{H(X)}{\ell\log_2(d)} \leq 1$

1er théorème de Shannon $\ell \geq H_d(X) = \dfrac{H(X)}{\log_2(d)}$

et égalité $\Leftrightarrow$ $\ell(x_i) = -\log_d(p(x_i))$

- **Codage par bloc** $Y = (X_1,\dots,X_s)$ $\Rightarrow$ $H_d(Y) \leq \ell_s < H_d(Y)+1$ et $H(Y) = sH(X)$

$\Rightarrow$ $H_d(X) \leq \ell < H_d(X) + \frac{1}{s}$

④ Algorithme de Compression

Rmq $p_i = p(x_i)$ et $l_i = l(x_i)$

- Condition d'optimalité ① $p_j > p_k \Rightarrow l_j \leqslant l_k$

② $l_N = l_{N-1}$

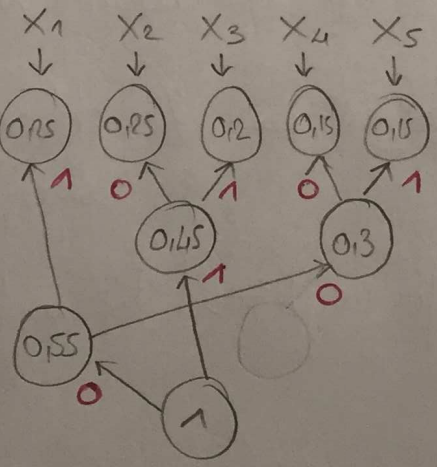
③ parmi les mots de longueur l_N, au moins deux ne diffère que par le dernier caractère

- Il est intéressant de coder les extensions de source
- Codage de Fano-Shannon principe du questionnaire ; On divise l'ensemble en deux sous ensembles $\simeq$ equiprobable.
- Codage de Huffmann On regroupe les deux évènements les moins probable pour créer un nouvel evenement

Ex Codage de Huffmann de $X = \{1,2,3,4,5\}$

de jeu de probabilités $\pi = (0{,}25; 0{,}25; 0{,}2; 0{,}15; 0{,}15)$

$\Rightarrow$ Arbre de Huffman



X	→ M	→ l
1	01	2
2	10	2
3	11	2
4	000	3
5	001	3

On a $l = 2{,}3$ et $H(X) = 2{,}28 \Rightarrow$ Code optimal

⑤ Canal & Capacités

$\mathcal{A}_N = \{x_1, \to x_N\}$ alphabet Entrée
$\mathcal{A}_M = \{y_1, \to y_M\}$ alphabet Sortie

• Chaine de l'information

Source → Codage Source → Codage Canal → Modulation → Canal → Demodulation
↓
Source estimée ← Décodage source ← Décodage canal

• **Canal** défini par $p(y_1, \to y_p ; x_1, \to x_p ; \text{état}) = \mathcal{P}$

• **Canal sans mémoire** $\mathcal{P} = p(y_1, x_1) \times \dots \times p(y_p, x_p)$

• **Matrice de Transition** $\underline{\underline{\Pi}} = [\Pi_{ij}] = [p(y_j | x_i)]_{\substack{i=1,\to N \\ j=1,\to M}}$

↳ $\underline{P_Y} = (P_Y^1, \to P_Y^M) = {}^t\underline{\underline{\Pi}}\,\underline{P_X}$

• **Canal uniforme % entrée** ⇔ Les symboles sont tous affectés de la même façon par les erreurs
⇔ Les lignes de $\underline{\underline{\Pi}}$ sont identiques à une permutation près

↳ Alors $H(Y|X) = H(Y|X=x_i)$ pour i∈Ω

• **Canal uniforme % sortie** ⇔ Les colonnes de $\underline{\underline{\Pi}}$ sont identiques à une permutation près

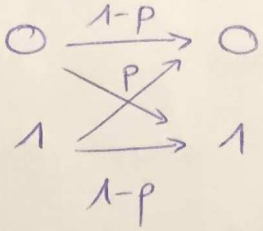
↳ Alors (P_X uniforme) ⇒ (P_Y uniforme)

• **Canal symétrique** ⇔ uniforme en entrée et sortie et $N = M$

- **Canaux à bruit additif gaussien** $y_k = x_k + b_k$ où $(b_k) \sim \mathcal{N}(0, \sigma^2)$

  Dès lors $p(y_k | x_k) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-|y_k - x_k|^2}{2\sigma^2}}$

- **Canal binaire symétrique**



D'où $\underline{\underline{\Pi}} = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$

- **Capacité d'un canal** $C = \max_{P_X} I(X, Y)$

  ↳ $C \leqslant \log_2(\min(M, N))$

  ↳ Pour un CS, $C = \log M + \sum_{j=1}^{M} p(y_j | x_i) \log p(y_j | x_i)$

- **Probabilité d'erreur** P_e → Proba de décoder x_i alors qu'on a envoyer x_j où $j \neq i$

Second théorème de Shanon

Lorsque $R = \frac{k}{n} < C$, il existe un code dont la P_e est aussi faible que l'on veut si $n \uparrow$

⑥ Codage Canal en Pratique

$\mathbb{F}_2 = (\{0, 1\}, \oplus, \cdot)$

- **Codage par bloc** On applique $C \begin{cases} \mathbb{F}_2^k \to \mathbb{F}_2^n \\ U \mapsto V \end{cases}$ où $n > k$

- **Codage linéaire par bloc** $\iff$ la répartition des mots codes formant un SEV de $\mathbb{F}_2^n$

- **Matrice génératrice** $G_{k \times n}$ tq $C(n) = nG$

- **G écrit de manière systématique** $G = [I_k \; P_{k, n-k}]$ (source ← I_k, redondance ← P_{k,n-k})

- **Matrice de contrôle de parité**

$$H \underset{(n-k)\times n}{\text{ tq }} G\,{}^tH = 0_{k\times(n-k)}$$

↳ Si G est sous forme systématique : $H = [\,{}^tP_{(n-k)\times k} \; I_{(n-k)(n-k)}\,]$

Rmq Les mots codes sont dans $\ker H$

- **Syndrome d'une séquence reçue r**

$$s(r) = r\,{}^tH \begin{cases} \neq 0 \Rightarrow \text{Une erreur s'est produite} \\ = 0 \Rightarrow r \text{ est un mot code mais pas forcément le bon} \end{cases}$$

↳ s est la somme des colonnes de H d'indices égaux aux positions des erreurs

- **Distance minimale** $d_{min} = \min_{i\neq j} d_H(c_i, c_j) = \min_{i\neq j} \sum_{k=1}^{n} c_{ik} \oplus c_{jk}$

- **Poid de Hamming** Si $z = (z_1, \dots, z_n)$ $P_z = \sum_{k=1}^{n} z_k$

↳ Dans le cas d'un codage linéaire par bloc

$$d_{min} = \min_{z\,(\text{mot code})} P_z$$

- **Capacité de Détection** $C_{det} = d_{min} - 1$

Correction $C_{corr} = \left\lfloor \frac{d_{min}-1}{2} \right\rfloor$