THEORIE

DE L'INFORMATION

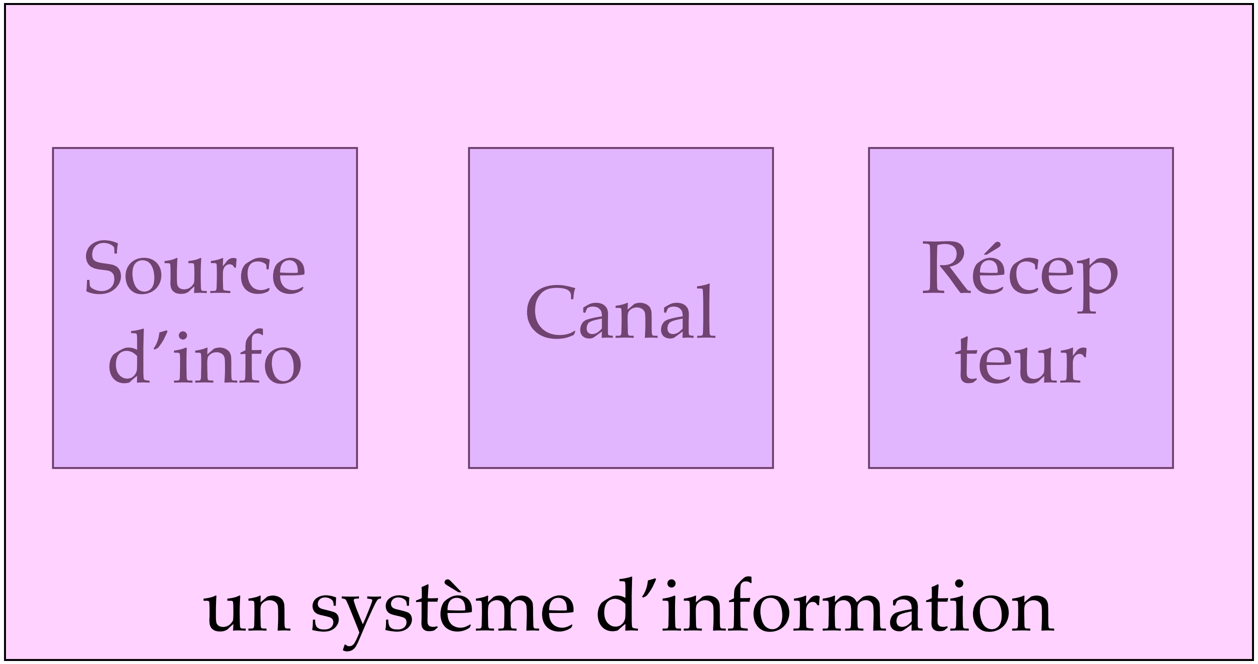
jp.guédon

INFO 3

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.1. Introduction**



3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- ## 3.2 Entropie de lois conjointes

Soit Z = (X,Y), la loi conjointe des 2 v.a. X et Y.

Dans le cours de probabilités on est intéressé de savoir comment les 2 v.a. se ressemblent : on mesure alors la corrélation entre les v.a.

Pour l'information, on est intéressé de savoir quelle partie d'information est portée par chacune des v.a. individuelles.

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.2 Entropie de lois conjointes**

Soit Z = (X,Y), la loi conjointe des 2 v.a. X et Y.

Définition

Si X possède n éléments ai et Y m éléments bj, on a :

$$H(Z) = H(X, Y) = E[-log2(p(X, Y))]$$

$$H(Z) = -\sum_i \sum_j [p_{ij} log2(p_{ij})] \quad \text{(bit/symbole)}$$

où p_{ij} est la probabilité p_{ij} = proba{(X=i) et (Y=j)}.

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

• 3.2 Entropie de lois conjointes

Si l'on écrit le théorème de Bayes :

$$proba((X=i)\,et\,(Y=j)) = proba(Y=j).\,proba((X=i)\,I\,(Y=j))$$

et

$$proba((X=i)\,et\,(Y=j)) = proba((X=i)).\,proba((Y=j)\,I\,(X=i))$$

en abrégé on a les formules : $p_{ij} = p_{i.p}(j\,I\,i) = p_{j.p}(i\,I\,j)$

$$H(Z) = -\sum_i \sum_j [p_{ij}\,log2(p_{ij})]$$

on réécrit alors H(Z) sous la forme :

$$H(Z) = -\sum_i \sum_j p_{ij}\,log2(p_i.\,p(j\,Ii))$$

cette écriture va nous permettre de faire les calculs après

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- ## 3.2 Entropie de lois conjointes

$$H(Z) = -\sum_{i}\sum_{j} p_{ij}\, log2\left(p_{i}\,.\,p(j \mid i)\right)$$

Source X	Canal	Récepteur Y

le système d'information Z

3. ENTROPIE D’UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

• 3.3 Entropie de lois conditionnelles

Le but ici est de calculer les relations
entre entropies de X, de Y et de Y∂X. Pourquoi ?
Lorsque nous appliquons tous ces résultats
à un système de communication, le système de messages X
envoie au travers du canal vers le destinataire Y :
les perturbations du canal seront ainsi modélisées par
les probabilités conditionnelles a posteriori.

H(X) est l'incertitude moyenne de X (quel que soit Y)
$H(X\partial Y) = H_Y(X)$ est l'incertitude moyenne de X
dans l'hypothèse ou Y est connu

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.3 Entropie de lois conditionnelles**

$$H(Z) = -\sum_i \sum_j p_{ij} \, log2\left(p_i \,.\, p(j\,I i)\right)$$

Source X	Canal	Récepteur Y

Entropie source :H(X),

Entropie source sachant destin Y: $H(X\partial Y) = H_Y(X)$

• 3.3 Entropie de lois conditionnelles

Calcul de H(X∂Y) = $H_Y(X)$

Décomposons le calcul en commençant par l'entropie de X sachant que Y=j :

$$H(X \, I \, j) = H(X \, I \, Y = j) = -\sum_i p(i \, I \, j) \, log2 \, p(i \, I \, j)$$

Puis on moyenne toutes les possibilités j sous la forme

$$H_Y(X) = H(X \, I \, Y) = E\left|H(X \, I \, Y = j)\right| = -\sum_j p(j) H(X \, I \, Y = j)$$

on obtient :

$$H_Y(X) = -\sum_i \sum_j p(j) \, p(i \, I \, j) \log_2(p(i \, I \, j))$$

$$H_Y(X) = -\sum_i \sum_j p_{ij} \log_2(p(i \, I \, j))$$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- ## 3.3 Entropie de lois conditionnelles

Calcul de $H(X\partial Y) = H_Y(X)$

$$H_Y(X) = -\sum_i \sum_j p_{ij} \log_2(p(i\,I\,j))$$

or $\log_2\{p(i\,\partial j)\} = \log_2\{p_{ij} / p_j\} = \log_2\{p_{ij}\} - \log_2\{p_j\}$,
donc

$$H_Y(X) = -\sum_i \sum_j p_{ij} \log_2(p_{ij}) + \sum_i \sum_j p_{ij} \log_2(p_j)$$

$$H_Y(X) = H(X,Y) + \sum_j \log_2(p_j) \sum_i p_{ij}$$

$$H_Y(X) = H(X,Y) + \sum_j \log_2(p_j)\, p_j$$

$$H_Y(X) = H(X,Y) - H(Y)$$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- ## 3.3 Entropie de lois conditionnelles

Calcul de $H(X\partial Y) = H_Y(X)$

$H_Y(X) = H(X,Y) — H(Y)$

De même

Calcul de $H(Y \partial X) = H_X(Y)$

$H_X(Y) = H(X,Y) — H(X)$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.3 Entropie de lois conditionnelles**

Calcul de $H(X\partial Y) = H_Y(X)$

$H_Y(X) + H(Y) = H(X,Y)$

De même

Calcul de $H(Y\partial X) = H_X(Y)$

$H_X(Y) + H(X) = H(X,Y)$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

• 3.3 Entropie de lois conditionnelles

Calcul de $H(X\partial Y) = H_Y(X)$

$H_Y(X) + H(Y) = H(X,Y)$

De même

Calcul de $H(Y\partial X) = H_X(Y)$

$H_X(Y) + H(X) = H(X,Y)$

$H(X) — H_Y(X) = H(Y) — H_X(Y)$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.3 Entropie de lois conditionnelles**

Calcul de $H(X \partial Y) = H_Y(X)$

$H_Y(X) + H(Y) = H(X,Y)$

De même

Calcul de $H(Y \partial X) = H_X(Y)$

$H_X(Y) + H(X) = H(X,Y)$

$H(X) — H_Y(X) = H(Y) — H_X(Y) \geq 0$

Car $H(X,Y)$ est une entropie

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

• 3.3 Entropie de lois conditionnelles

$$H(X) - H_Y(X) = H(Y) - H_X(Y) \geq 0$$

Car H(X,Y) est une entropie

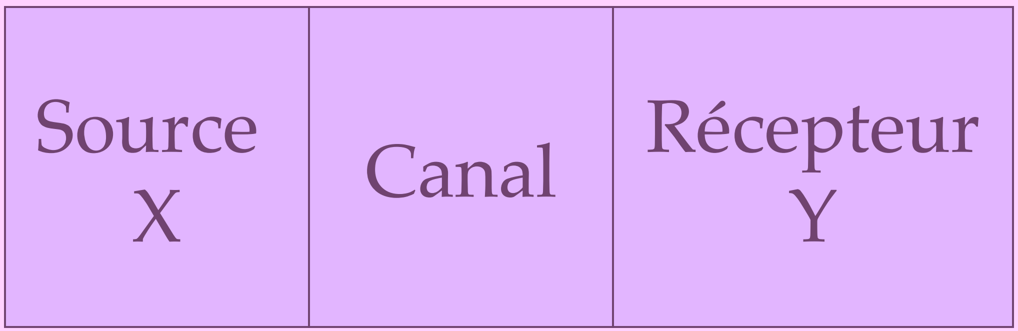
Le désordre est plus grand sur la v.a. X que sur cette v.a. lorsque l'on ajoute la connaissance de Y

Si X et Y sont indépendantes alors $H_Y(X) = H(X)$
(l'info apportée par Y ne sert à rien pour X)

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.4 de l'entropie à l'information mutuelle**

$$H_Y(X) = -\sum_i \sum_j p_{ij} \log_2\left(p(i \mid j)\right)$$



$$H_Y(X) = H(X,Y) - H(Y)$$

$H_Y(X)$: Incertitude moyenne de X sachant Y

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.4 de l'entropie à l'information mutuelle**

On veut mesurer l'info de X transmise à Y

Source X	Canal	Récepteur Y

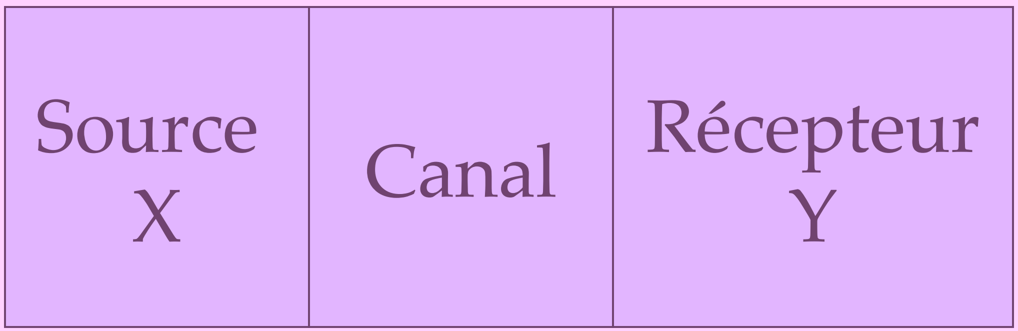
$$H_Y(X) = H(X,Y) - H(Y)$$

$H_Y(X)$: Incertitude moyenne de X sachant Y

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.4 de l'entropie à l'information mutuelle**

On veut mesurer l'info de X transmise à Y



Cas canal parfait : tout est transmis
Cas canal bouché : rien ne passe

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.4 de l'entropie à l'information mutuelle**

On veut mesurer l'info de X transmise à Y

On définit l'information mutuelle:

$$I(X,Y) = H(X) - H_Y(X)$$

I représente l'information contenue dans la v.a. Y au sujet de la v.a. X

$$I(X,Y) \geq 0$$

Car $H(X) — H_Y(X) \geq 0$

$H_Y(X)$: Incertitude moyenne de X sachant Y

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.4 de l'entropie à l'information mutuelle**

On définit l'information mutuelle:

$$I(X,Y) = H(X) - H_Y(X)$$

I : information de Y au sujet de X

$H(X)$: info de la source

$H_Y(X)$: « équivoque » : perte d'info dans le canal

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.4 de l'entropie à l'information mutuelle**

$$H(X) - H_Y(X) = I(X,Y)$$

Source X	Canal	Récepteur Y

$I(X,Y)$: info reçue par destination

$H(X)$: info envoyée de la source

$H_Y(X)$: perte d'info dans le canal

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.4 de l'entropie à l'information mutuelle**

$$I(X,Y) = H(X) - H_Y(X)$$

$$I(X,Y) = H(X) + H(Y) — H(X,Y)$$

$$I(X,Y) = I(Y,X)$$

i.e. $I(X,Y) = H(X) — H_Y(X) = H(Y) — H_X(Y)$.

soit

$$I(X,Y) = -\sum_i \sum_j p_{ij} \log_2\left(\frac{p_{ij}}{p_i p_j}\right)$$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.4 de l'entropie à l'information mutuelle**

Les 2 cas extrêmes :

Si X et Y sont indépendantes Alors $I(X,Y) = 0$
car $I(X,Y) = H(X) — H_Y(X)$

Si X et Y sont liées $(H(Y)=H(X)=H(X,Y))$
Alors $I(X,Y) = H(X)$
car $I(X,Y) = H(X) + H(Y) — H(X,Y)$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.5 le canal binaire symétrique (CBS)**

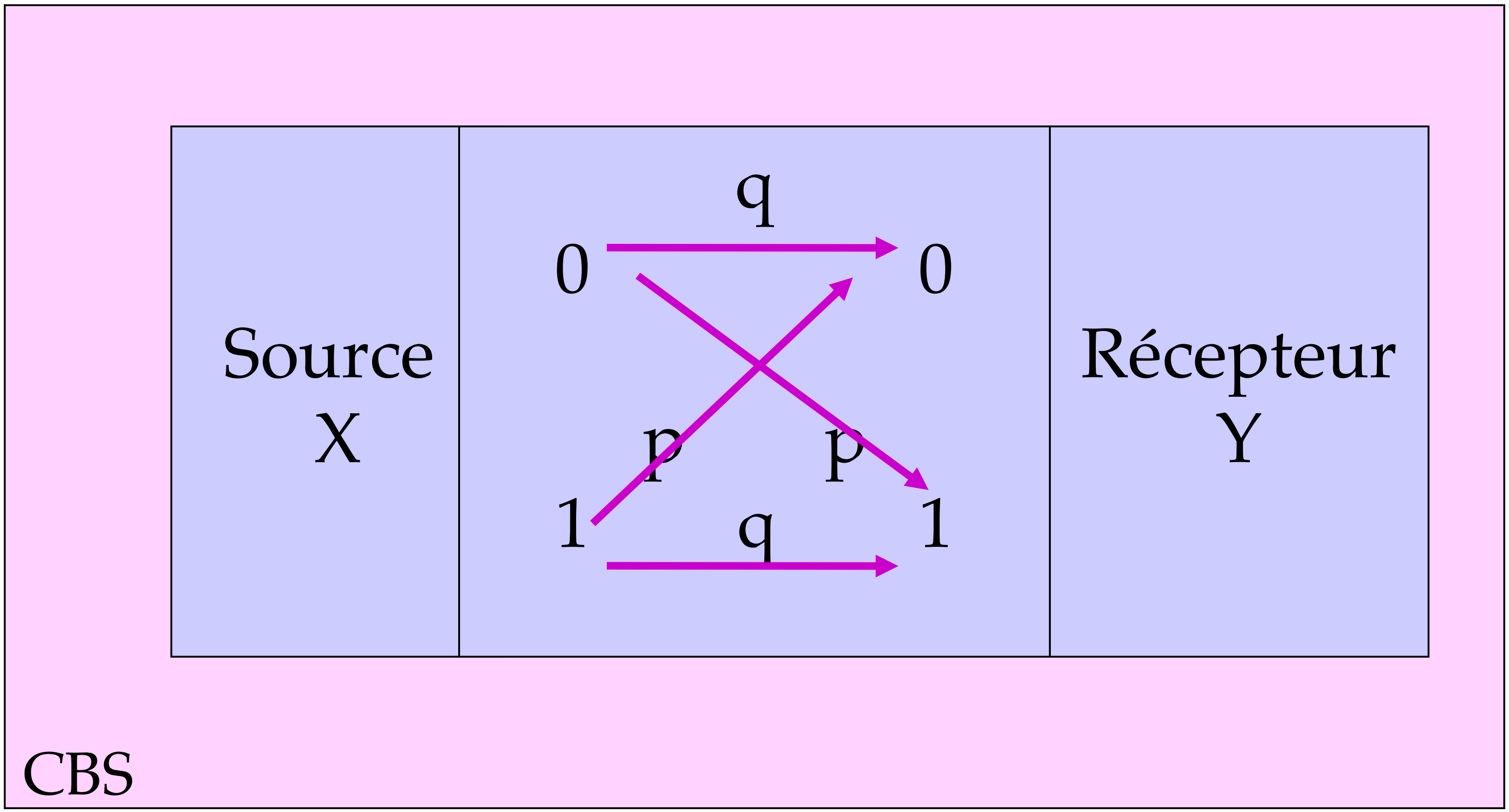
On considère le cas le plus simple.

La source envoie des bits d'information,
qui sont entachés d'erreur au niveau du canal
et sont reçus par la destination.
On suppose de plus que l'erreur entachant un bit
est identique que la valeur de ce bit soit 0 ou bien 1.

On appelle cela "le canal binaire symétrique"
ou CBS.

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.5 le canal binaire symétrique (CBS)**



- **3.5 le canal binaire symétrique (CBS)**

Écrivons les équations :

$$\Pr\{Y=0 \, \partial X = 0\} = q = \Pr\{Y=1 \, \partial X = 1\}$$

$$\Pr\{Y=0 \, \partial X = 1\} = p = \Pr\{Y=1 \, \partial X = 0\}$$

$$p + q = 1$$

q est la probabilité de bonne transmission
et p la probabilité d'erreur sur un bit

Nous sommes en mesure de calculer les entropies
et donc l'information mutuelle de ce système :
on a $p_j(i) = \Pr\{Y = i \, \partial X = j\}$

- **3.5 le canal binaire symétrique (CBS)**

$$H(Y\ I\ X=0) = -\sum_{i} p_j\, log2(p_j)$$

$$H(Y\ I\ X=0) = -(q \log_2(q) + p \log_2(p))$$

De même $$H(Y\ I\ X=1) = -(p \log_2(p) + q \log_2(q))$$

Puisque q = p—1 on écrit $H(Y\partial X=0) = H(Y\partial X=1) = H(p)$

On peut donc maintenant calculer

$H_X(Y) = H(Y\partial X) = E\{H(Y\partial X= ai)\} = E(constante) = H(p)$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.5 le canal binaire symétrique (CBS)**

donc $H_X(Y) = H(p)$

Il nous reste à exprimer H(Y) pour en déduire I

$I(X,Y) = H(Y) — H_X(Y).$

On a

$\Pr\{Y = 0\} = \Pr\{X=0\}.\Pr\{Y=0 \,\partial X=0\} + \Pr\{X=1\}.\Pr\{Y=0 \,\partial X=1\}$

$\Pr\{Y = 0\} = \Pr\{X=0\}.\ q + \Pr\{X=1\}.p$

$\Pr\{Y = 0\} = (1—p)\ P_0 + p\ (1—P_0) = P_0 + p —2pP_0 = r$

de même

$\Pr\{Y = 1\} = \Pr\{X=1\}.\ q + \Pr\{X=0\}.p$

$\Pr\{Y = 1\} = (1—p)(1—P_0) + pP_0$

$\Pr\{Y = 1\} = 1—p —P_0 + 2pP_0 = 1 —\Pr\{Y = 0\} = 1 —r$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.5 le canal binaire symétrique (CBS)**

C'est à dire que l'on ne connaît pas la distribution
de probabilité de la source X
(on dit $\Pr\{X=0\}=P_0$ et $\Pr\{X=1\} = 1 - P_0$),
ni de celle de Y

MAIS que l'on peut les modéliser sous
forme de loi de Bernoulli car on n'a que les 2 seuls états
si X suit Bernoulli (p,q) alors Y suit Bernoulli (r, 1 —r).
Finalement : $H(Y) = - \{r \log_2 r + (1-r) \log_2(1-r)\}$
$H(Y)= H(r) = H(P_0 + p - 2pP_0)$

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.5 le canal binaire symétrique (CBS)**

En résumé :

$H_X(Y) = H(p) = —(p \log_2 p + q \log_2 q))$

et

$H(Y) = — \{r \log_2 r + (1—r) \log_2(1—r)\}$

$H(Y) = H(r) = H(P_0 + p —2pP_0)$

L'information transmise par le canal est donc :

$$I(X,Y) = H(r) — H(p) = H(P_0 + p —2pP_0) — H(p)$$

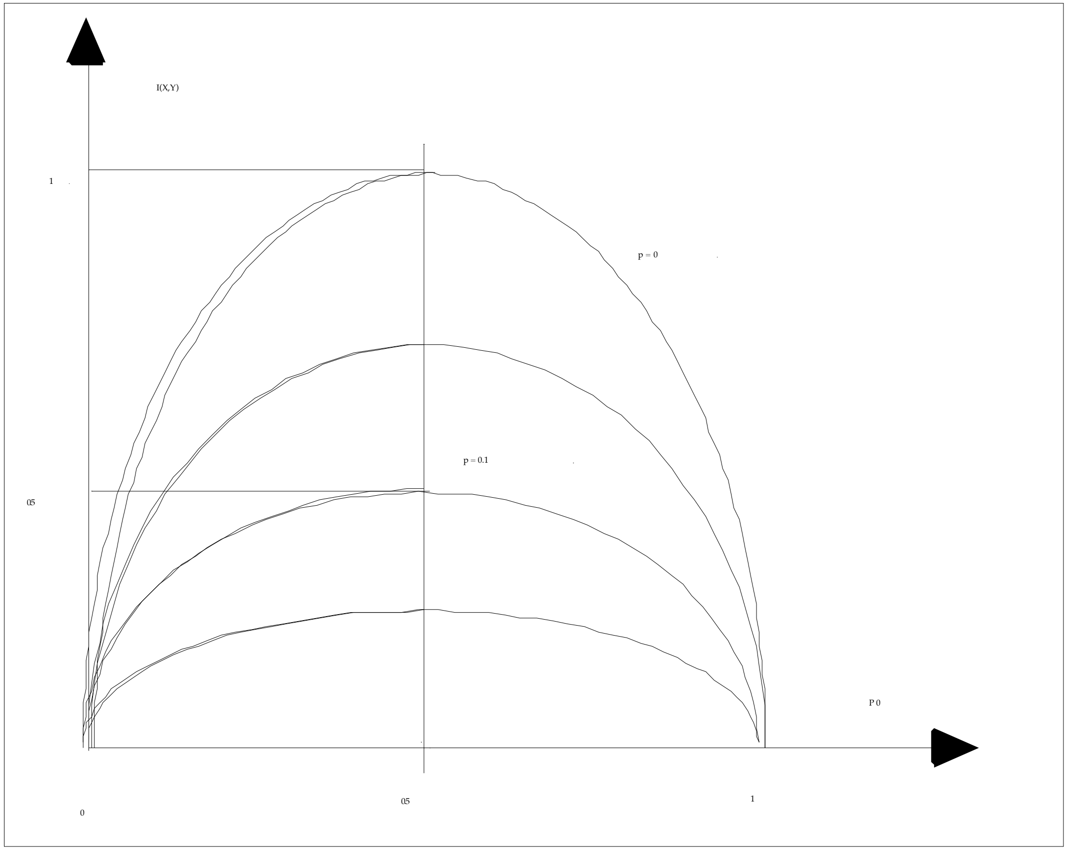
cela dépend donc à la fois du canal (la valeur de p) et de la distribution initiale de la source (la valeur de P_0) .

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.5 le canal binaire symétrique (CBS)**

Si l'on veut optimiser la transmission dans ce canal
(optimum de I(X,Y))
il faut donc optimiser d'une part la répartition
de la loi de probabilité de X
D'autre part le canal

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL



tracé de I(X,Y) selon X (valeur de p)
et le canal (valeur de P0)

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

La conclusion est que pour
$P_0 = 0.5$ (source dite symétrique)
la transmission sera idéale
quelque soit la dégradation du
canal. On peut noter que lorsque
P_0 est proche de 0 ou de 1
(la source n'envoie presque que des 1 ou des 0)
la quantité d'information transmise par le canal s'effondre
quel que soient ses caractéristiques intrinsèques
(radioélectrique, cuivre, optique).
Cela provient simplement de la nature de la source.

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

Pour $P_0 = 0.5$

on obtient $r = 0.5$

et $I(X,Y) = H(0.5) — H(p)$

canal sans bruit (p = 0)	$I = H(0.5) = — 2(0.5 \log_2 0.5)$ I = 1 bit par digit 0/1 envoyé
Canal optique p varie de 10^{-12} à 10^{-6}	I est presque égal à 1
Canal hertzien $p = 10^{-3}$	
à $p = 10^{-2}$	$I = 0.992$
$p = 10^{-1}$	$I = 0.5$
$p = 0.5$	I = 0 : la capacité du canal est nulle

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.6 la capacité du canal**

L'optimum de I(X,Y) est appelée la Capacité

$$C = Max\left(H\left(X\right) - Hy\left(X\right)\right)$$

- 3.7 Theorème de Shannon

Théorème de Shannon (sur codage et capacité du canal)
Soit un canal de capacité C et une source d'entropie H
C = Max (H(X) - Hy(X))
Si H <= C il existe un codage de la source pour la transmettre sur le canal avec un taux d'erreur aussi petit que l'on désire
Si H >C il n'existe aucun codage qui ne donne une Équivoque inférieure à H - C

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.7 Theorème de Shannon**

Soit un canal de capacité C et une source d'entropie H

$$C = \underset{Px}{\text{Max}} \left(H(X) - Hy(X) \right) = \underset{Px}{\max} \{ I(Xi,Yi) \}$$

La capacité C est définie uniquement à partir de l'information mutuelle

Un bon système de codage (vu du coté du canal) est donc constitué d'un code qui permet de se rapprocher de la valeur C autant que faire se peut.

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.7 Theorème de Shannon**

Soit un canal de capacité C et une source d'entropie H

C = Max (H(X) - Hy(X)) = max{ I(Xi,Yi)}

Px Px

Exemple: capacité d'un canal binaire symétrique

Pr{yi∂xi} = q=1-p si yi=xi et p sinon

Shannon:

C = max{ I(Xi,Yi)} = max { H(X) - Hy(X) }

Px Px

3. ENTROPIE D'UN SYSTEME INFORMATION CONJOINTE, CAPACITE DU CANAL

- **3.7 Theorème de Shannon**

Exemple: capacité d'un canal binaire symétrique

On sait que

$I(X,Y) = H(r) - H(p) = H(P_0 + p - 2pP_0) - H(p)$

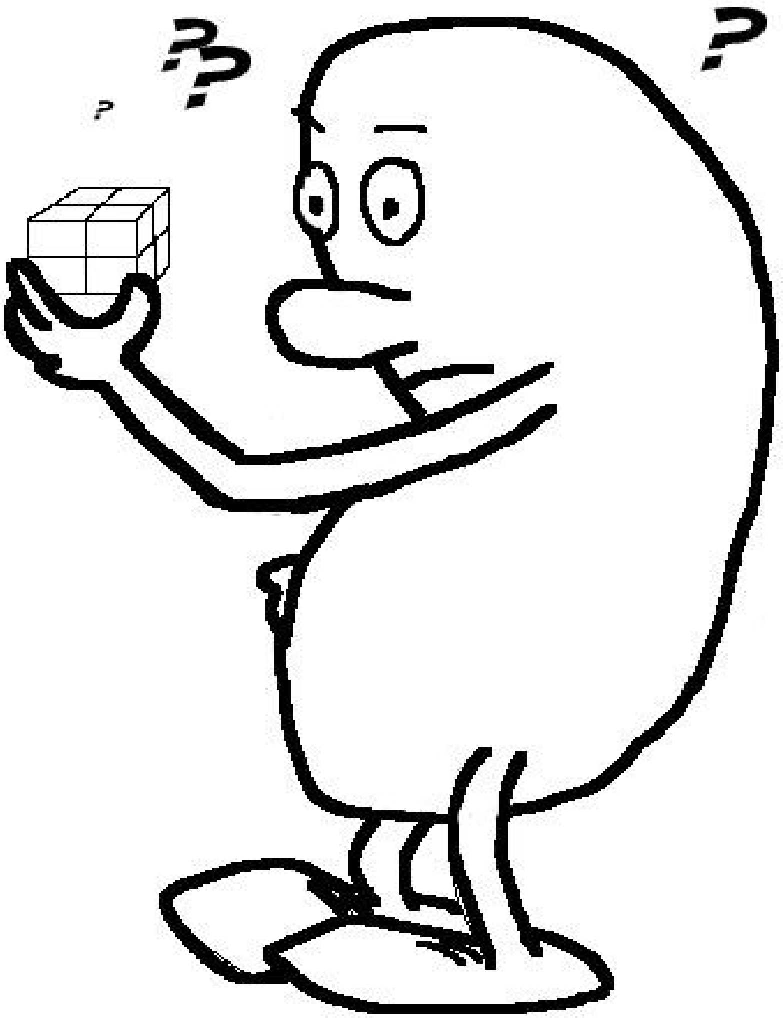
Avec $\Pr\{Y=0 \partial X = 0\} = 1\text{-}p = \Pr\{Y=1 \partial X = 1\}$

Et $\Pr\{Y=0 \partial X = 1\} = p = \Pr\{Y=1 \partial X = 0\}$.

Le maximum des P_X est obtenu pour $P_0 = 0.5$:

$C = 1 - H(p) = 1 - (p \log_2 p) + (1\text{-}p) \log_2 (1\text{-}p))$

Théorie de l’ Information



- **merci**