Réseaux, information et communications (INFO-F303) Partie Théorie de l'Information 5. Canal bruité

Christophe Petit

Université libre de Bruxelles

Plan du cours

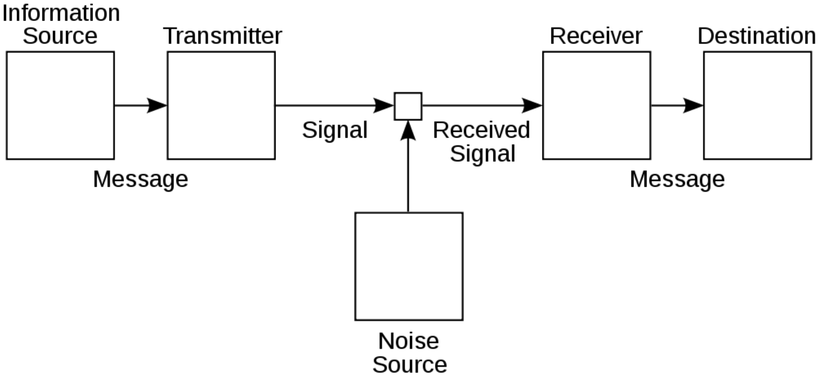
1. Notion de code
2. Source aléatoire et codes efficaces
3. Entropie et codage efficace
4. Compression sans perte
5. Canal bruité
6. Codes correcteurs d'erreurs
7. Codes linéaires
8. Quelques familles de codes linéaires

A. Rappels mathématiques (chapitre 7.1 du syllabus)

Canal bruité

- Buts :
  - modéliser un processus de transmission d'information avec erreurs
  - quantifier l'information transmise ou perdue
- Canal markovien
- Entropie résiduelle, conditionnelle, croisée ; information mutuelle
- Capacité d'un canal (quantité d'information maximale qu'il peut transporter de l'émetteur au récepteur)
- Cas du canal symétrique

Canaux et bruit



Canal de communication

- **Emetteur** envoie un message $x_i \in X$, avec $|X| = m$
- **Récepteur** reçoit $y_i \in Y$, avec $|Y| = n \geq m$
- Erreurs de transmission dues au **canal bruité**

Probabilités conditionnelle, conjointe, marginale

- **Probabilité conditionnelle** $\mathbb{P}[y_j|x_i]$
  probabilité de recevoir y_j sachant que x_i a été envoyé
- On a
$$\forall i \;:\; \sum_{j=1}^{n} \mathbb{P}[y_j|x_i] = 1$$
- **Probabilité conjointe ou liée**
$$\mathbb{P}[x_i, y_j] = \mathbb{P}[x_i] \cdot \mathbb{P}[y_j|x_i]$$
- **Probabilité marginale**
$$\mathbb{P}[y_j] = \sum_{i=1}^{m} \mathbb{P}[x_i, y_j]$$

Questions

- Etant donné un canal bruité caractérisé par des probabilités conditionnelles $\mathbb{P}[y_j|x_i]$
- A quel point une valeur source détermine-t-elle une valeur observée ?
- A quel point une valeur observée détermine-t-elle une valeur source ?
- Comment quantifier cette incertitude / information ?
- Comment émettre à la source pour optimiser l'information transmise par le canal bruité ?

Entropie résiduelle

- Supposons que récepteur reçoive y_j
- L'incertitude sur la valeur x_i envoyée est appelée **entropie résiduelle**

$$\forall j : \; H_b(X|y_j) = -\sum_{i=1}^{m} \mathbb{P}[x_i|y_j] \cdot \log_b \mathbb{P}[x_i|y_j]$$

Entropie résiduelle : interprétation

- Si $H_b(X|y_j) = 0$: incertitude nulle
  - Valeur x_i envoyée déterminée par y_j
  - $\mathbb{P}[x_i|y_j] = 1$ pour le x_i correspondant, et 0 ailleurs

- Si $H_b(X|y_j) > 0$
  - Connaissance de y_j pas suffisante pour déterminer la valeur x_i envoyée par Alice
  - On aurait besoin, en moyenne, de $H_b(X|y_j) \cdot \log_2 b$ bits supplémentaires pour déterminer x_i

(remarque : base b implicite dans la suite)

Entropie conditionnelle

- **Valeur moyenne de l'entropie résiduelle** sur toutes les valeurs possibles de y_j

$$H(X|Y) = \sum_{j=1}^{n} \mathbb{P}[y_j] \cdot H(X|y_j)$$

- On déduit

$$H(X|Y) = -\sum_{i,j} \mathbb{P}[x_i, y_j] \cdot \log \mathbb{P}[x_i|y_j]$$

Information mutuelle

- On a

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

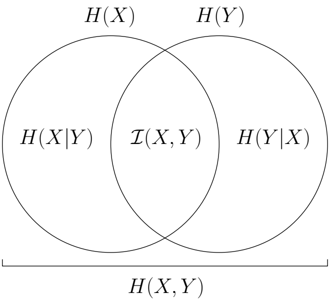
- **Information mutuelle** de X et Y est

$$\mathcal{I}(X, Y) := H(X, Y) - H(X|Y) - H(Y|X)$$

- On déduit

$$\begin{aligned} \mathcal{I}(X, Y) &= H(X) - H(X|Y) = H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X, Y) \end{aligned}$$

Relations entre quantités d'information



- X et Y corrélées, pas nécessairement indépendantes donc $H(X,Y) \leq H(X) + H(Y)$
- $H(X|Y)$ entropie conditionelle (incertitude sur X quand Y est fixée, en moyenne)
- $\mathcal{I}(X,Y)$ information mutuelle, simultanément dans les deux variables, transmise par le canal

Information mutuelle : forme close

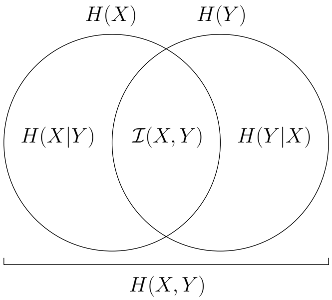
- On a

$$\begin{aligned}
\mathcal{I}(X,Y) &= H(X,Y) - H(X|Y) - H(Y|X) \\
&= \sum_{i,j} \mathbb{P}[x_i,y_j](-\log \mathbb{P}[x_i,y_j] + \log \mathbb{P}[x_i|y_j] + \log \mathbb{P}[y_j|x_i]) \\
&= \sum_{i,j} \mathbb{P}[x_i,y_j](-\log \mathbb{P}[y_j] + \log \mathbb{P}[y_j|x_i]) \\
&= -\sum_{i,j} \mathbb{P}[x_i,y_j] \log \frac{\mathbb{P}[y_j]}{\mathbb{P}[y_j|x_i]} \\
&= -\sum_{i,j} \mathbb{P}[x_i,y_j] \log \frac{\mathbb{P}[x_i]\mathbb{P}[y_j]}{\mathbb{P}[x_i,y_j]}
\end{aligned}$$

L'information mutuelle est positive

- On a $\mathcal{I}(X,Y) = -\sum_{i,j} \mathbb{P}[x_i, y_j] \log \frac{\mathbb{P}[x_i]\mathbb{P}[y_j]}{\mathbb{P}[x_i, y_j]}$
- Et donc $\mathcal{I}(X,Y) \geq 0$ par l'inégalité de Gibbs (cfr. chapitre 3)
- On a $\mathcal{I}(X,Y) = 0 \Leftrightarrow \mathbb{P}[x_i, y_j] = \mathbb{P}[x_i]\mathbb{P}[y_j]$ $\Leftrightarrow X$ et Y sont indépendantes
- Par contre $-\log \frac{\mathbb{P}[x_i]\mathbb{P}[y_j]}{\mathbb{P}[x_i, y_j]}$ la **différence d'information mutuelle** pour le couple (x_i, y_j), peut être négative (si corrélation négative $\mathbb{P}[x_i, y_j] \leq \mathbb{P}[x_i]\mathbb{P}[y_j]$)

Maximiser $\mathcal{I}(X, Y)$?



$\mathcal{I}(X, Y)$ information mutuelle, simultanément dans les deux variables, transmise par le canal

Capacité d'un canal

- Information mutuelle entre messages émis X et reçu Y dépend des probabilités de transmission du canal $\mathbb{P}[y_j|x_i]$ mais aussi de la distribution de X
- **Capacité** d'un canal de communication est le **maximum de l'information mutuelle**, sur l'ensemble des distributions pour X

$$\mathcal{C} = \max_{\mathbb{P}[X]} \mathcal{I}(X, Y)$$

- Bien définie, et ne dépend que des probabilités $\mathbb{P}[y_j|x_i]$

Calcul de la capacité

- Soit $p_{ij} = \mathbb{P}[y_j|x_i]$ constantes pour un canal donné
- Problème d'optimisation non linéaire : choisir $p_i = \mathbb{P}[x_i]$ tels que $0 \leq p_i \leq 1$, $\sum_i p_i = 1$ et $\mathcal{I}(X, Y)$ maximal
- Calcul analytique possible dans certains cas (par exemple, canal symétrique)
- On peut montrer : lorsque $\mathcal{I}(X, Y)$ est maximale, toutes les *informations mutuelles conditionnelles*

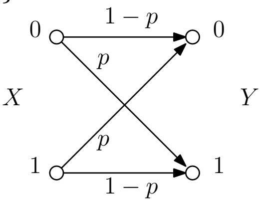
$$\mathcal{I}(Y|x_i) = -\sum_j \mathbb{P}[y_j|x_i] \log \frac{\mathbb{P}[y_j]}{\mathbb{P}[y_j|x_i]}$$

  sont égales à la capacité

Exemple : canal binaire symétrique

- X et Y ont des valeurs dans $\{0, 1\}$

$$\mathbb{P}[y_j|x_i] = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$



- Soit $p_0 = \mathbb{P}[X = 0]$. On a

$$\mathbb{P}[x_i, y_j] = \begin{bmatrix} p_0\,(1-p) & p_0\,p \\ (1-p_0)\,p & (1-p_0)(1-p) \end{bmatrix}$$

Capacité du canal binaire symétrique

- Notons $\mathcal{H}(p) = -p\log_2 p - (1-p)\log_2(1-p)$
- $$\begin{aligned} H(Y|X) = & -p_0(1-p)\log(1-p) - p_0 p \log p \\ & -(1-p_0)p\log p - (1-p_0)(1-p)\log(1-p) \\ = & \ \mathcal{H}(p) \end{aligned}$$
- Soit $q_0 = \mathbb{P}[Y=0] = p_0(1-p) + (1-p_0)p$
- $H(Y) = -q_0\log q_0 - (1-q_0)\log(1-q_0) = \mathcal{H}(q_0)$
- $\mathcal{I}(X,Y) = H(Y) - H(Y|X) = \mathcal{H}(q_0) - \mathcal{H}(p)$
- $\mathcal{I}(X,Y)$ maximal $\Leftrightarrow \mathcal{H}(q_0)$ maximal $\Leftrightarrow q_0 = \frac{1}{2} \Leftrightarrow p_0 = \frac{1}{2}$
- $\mathcal{C} = 1 - \mathcal{H}(p)$
- $\mathcal{C} = 0$ si $p = \frac{1}{2}$; $\qquad \mathcal{C} = 1$ si $p \in \{0,1\}$

Exemple : canal symétrique général

- r symboles ; probabilité d'erreur p équidistribuée sur les autres symboles

$$\mathbb{P}[y_j|x_i] = \begin{cases} 1-p & \text{si} \quad i=j \\ \frac{p}{r-1} & \text{sinon} \end{cases}$$

- $H_r(Y|x_i) = -(1-p)\log_r(1-p) - (r-1)\frac{p}{r-1}\log_r\frac{p}{r-1}$ $= \mathcal{H}_r(p)$
- Donc $H_r(Y|X) = \mathcal{H}_r(p)$
- $\mathcal{I}_r(X,Y) = H_r(Y) - H_r(Y|X) = H_r(Y) - \mathcal{H}_r(p)$
- Pour une source uniforme $\mathcal{C}_r = 1 - \mathcal{H}_r(p)$
- $\mathcal{C}_r = 0$ si $p = \frac{r-1}{r}$; $\quad \mathcal{C} = 1$ si $p \in \{0,1\}$

Questions ?

?

Crédits et remerciements

- Mes transparents suivent fortement les notes de cours développées par le Professeur Yves Roggeman pour le cours INFO-F303 à l'Université libre de Bruxelles
- Une partie des transparents et des exercices ont été repris ou adaptés des transparents développés par le Professeur Jean Cardinal pour ce même cours
- Je remercie chaleureusement Yves et Jean pour la mise à disposition de ce matériel pédagogique, et de manière plus large pour toute l'aide apportée pour la reprise de ce cours
- Les typos et erreurs sont exclusivement miennes (merci de les signaler !)