*Nicolas Sendrier*

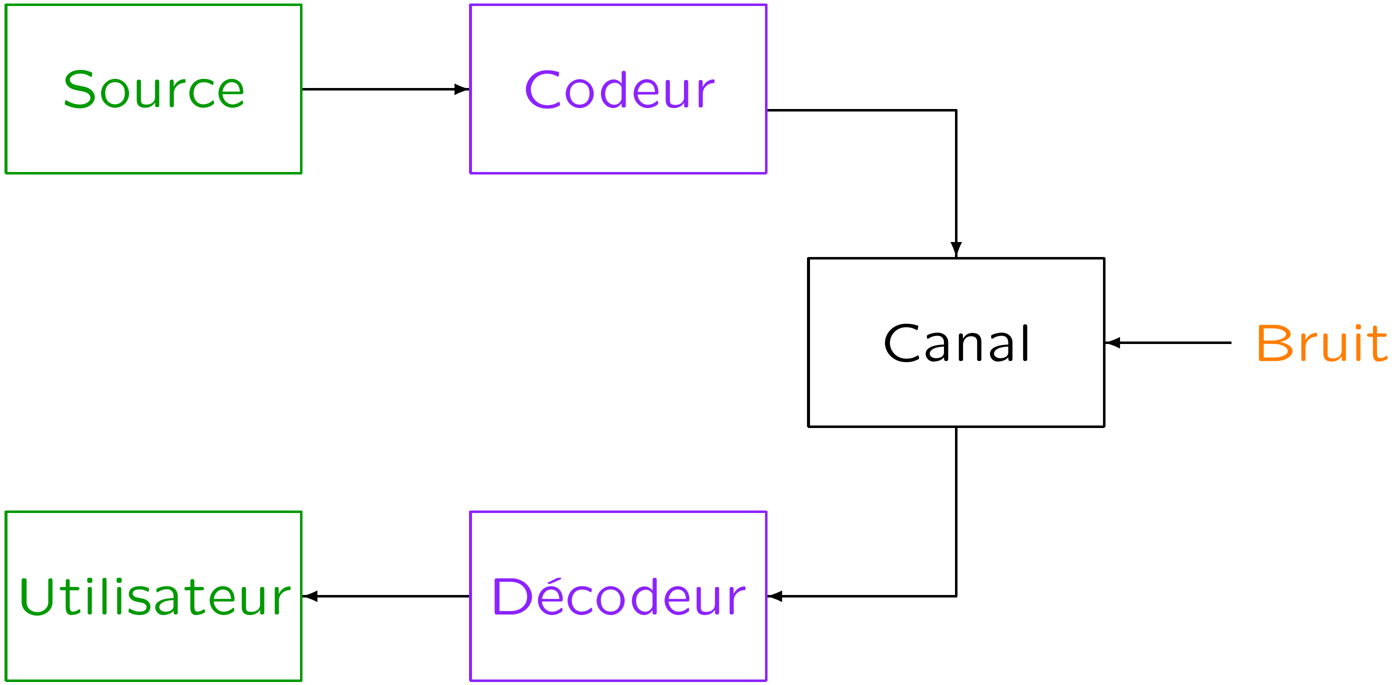
## Majeure d'informatique

# Introduction  la théorie de l'information

Cours n°0

**Présentation**

# Système de communication

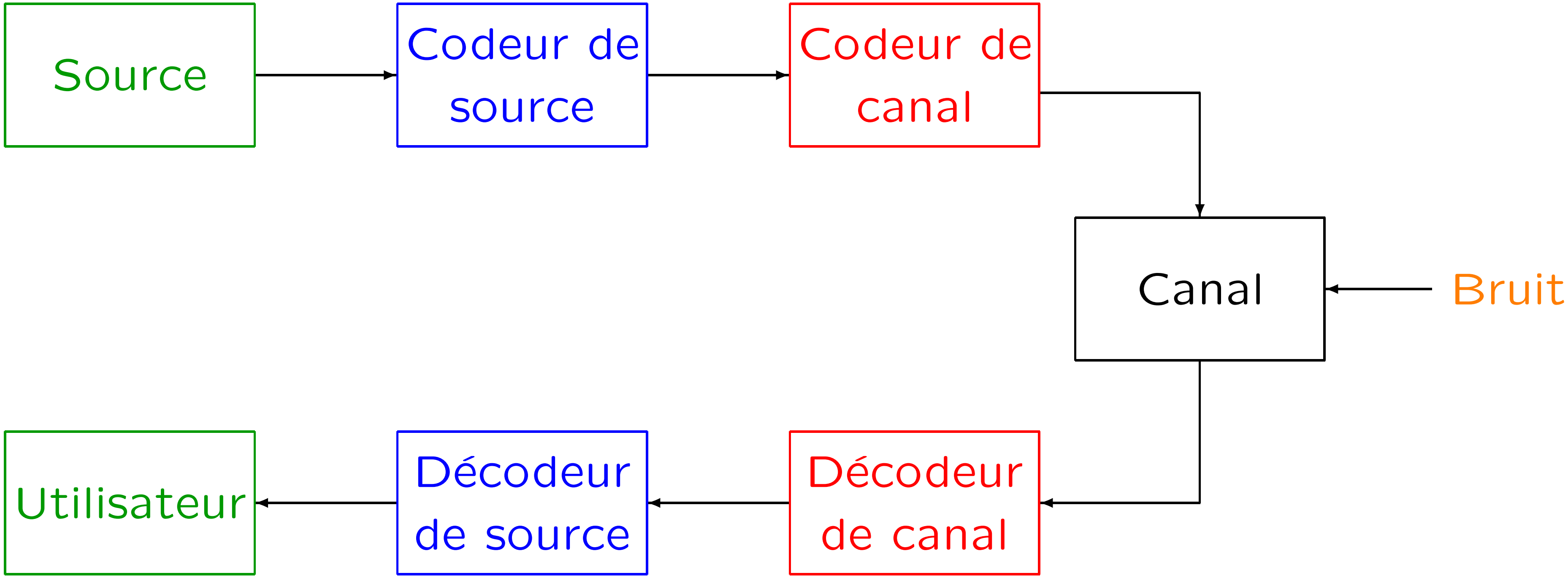



**Source :** voix, musique, image (fixe ou animée), texte, . . .

**Canal :** radio, fil, fibre optique, support magnétique ou optique, . . .

**Bruit :** perturbations electromagnétiques, rayures, . . .

# Codage de source et de canal




**Efficacité :** Pour faire parvenir une quantité donnée d'information à l'utilisateur, utiliser le minimum de ressources.

**Fiabilité :** Restituer à l'utilisateur une information *suffisamment* fidèle à celle produite par la source.

## Exemple de codage de source

$X = \{a_1, a_2, a_3, a_4\}$, Loi 1 : $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$, Loi 2 : $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right)$

Code A : $\begin{cases} a_1 \longrightarrow 00 \\ a_2 \longrightarrow 01 \\ a_3 \longrightarrow 10 \\ a_4 \longrightarrow 11 \end{cases}$ Code B : $\begin{cases} a_1 \longrightarrow 0 \\ a_2 \longrightarrow 10 \\ a_3 \longrightarrow 110 \\ a_4 \longrightarrow 111 \end{cases}$

**Longueur moyenne :**

- pour le Code A, on trouve 2 dans les deux cas
- pour le Code B
  - avec la Loi 1 : $(1+2+3+3) \times \frac{1}{4} = \frac{9}{4} = 2.25$
  - avec la Loi 2 : $1 \times \frac{1}{2} + 2 \times \frac{1}{4} + (3+3) \times \frac{1}{8} = \frac{7}{4} = 1.75$

*Le meilleur code dépend de la loi d'émission de la source*

# Entropie d'une source discrète

Source discrète sans mémoire :

– Alphabet fini $X = \{a_1, \ldots, a_K\}$

– Loi de probabilité $P_X(a_1), \ldots, P_X(a_K)$

La **longueur moyenne** de tout code est au moins égale à l'**entropie** :

$$H(X) = \sum_{k=1}^{K} P_X(a_k) \log_2 \frac{1}{P_X(a_k)}$$

Dans l'exemple précédent

– Loi 1 : $H(X) = 4 \times \left(\frac{1}{4} \times 2\right) = 2$

– Loi 2 : $H(X) = \frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{8} \times 3 = \frac{7}{4} = 1.75$

## Un autre exemple de codage de source

$X = \{a_1, a_2, a_3, a_4\}$, Loi : $(0.4, 0.3, 0.2, 0.1)$

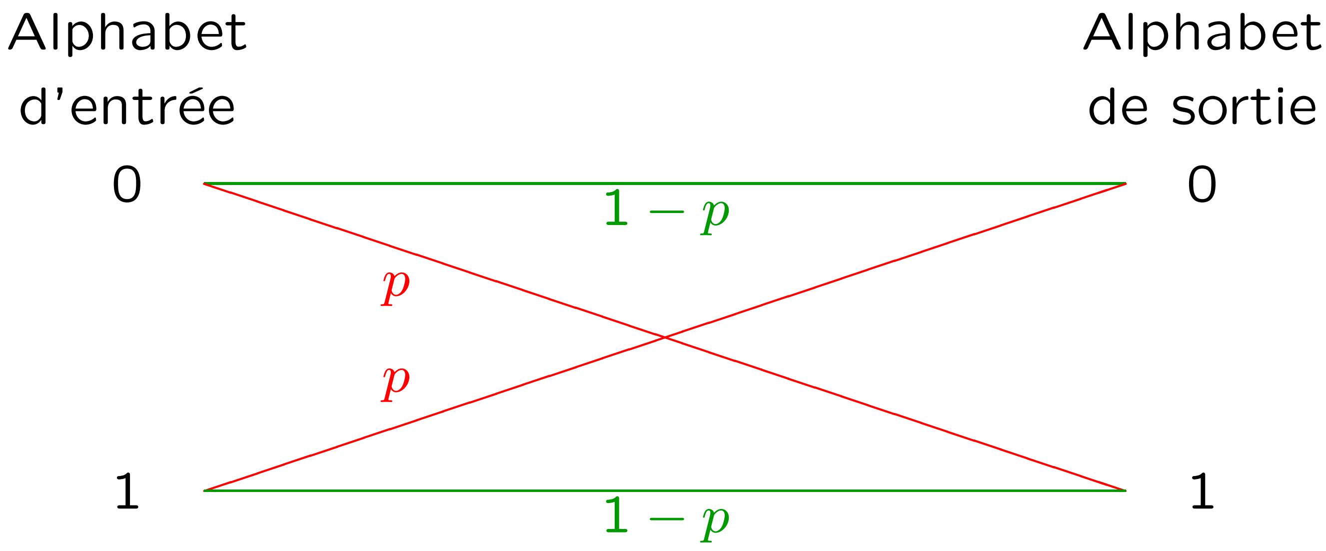
$$\text{Code optimal : } \begin{cases} a_1 \longrightarrow 0 \\ a_2 \longrightarrow 10 \\ a_3 \longrightarrow 110 \\ a_4 \longrightarrow 111 \end{cases}$$

**Entropie :** 1.846

**Longueur moyenne :** 1.9

Le code optimal n'atteint l'entropie que si les probabilités sont des puissances de $\frac{1}{2}$.

# Canal binaire symétrique




$p$ est la probabilité d'erreur du canal.

Pour combattre les effets du bruit on ajoutera de la redondance. Par exemple, le code à répétition de longueur 3 :

$$
\begin{array}{rcl}
0 & \mapsto & 000 \\
1 & \mapsto & 111
\end{array}
$$

Ce code à un taux de transmission 0.33 et corrige une erreur.

# Code à répétition

Si la probabilité d'erreur du canal $p = 0.01$ pour chaque symbole transmis, il se produira 0 ou 1 erreur avec une probabilité

$$(1-p)^3 + 3p(1-p)^2 \approx 0.9997$$

et il se produira 2 ou 3 erreurs avec une probabilité

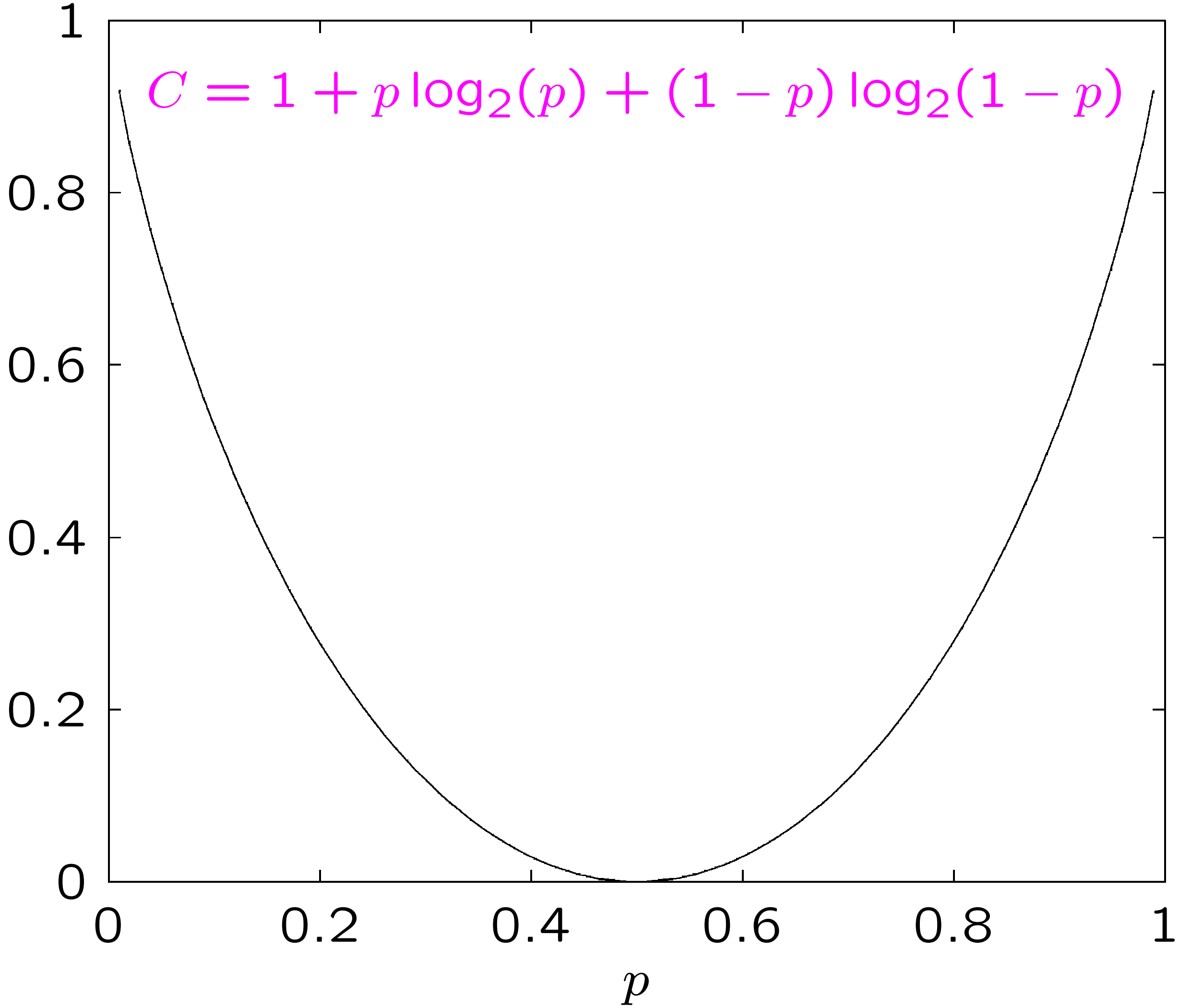
$$3p^2(1-p) + p^3 \approx 3 \times 10^{-4}$$

Le symbole sera mal transmis avec une probabilité $3 \times 10^{-4}$. Avec un code à répétition de longueur 5, cette probabilité tombe à $10^{-5}$

$$10p^3(1-p)^2 + 5p^4(1-p) + p^5 \approx 10^{-5}$$

Ce code à un taux de transmission 0.2.

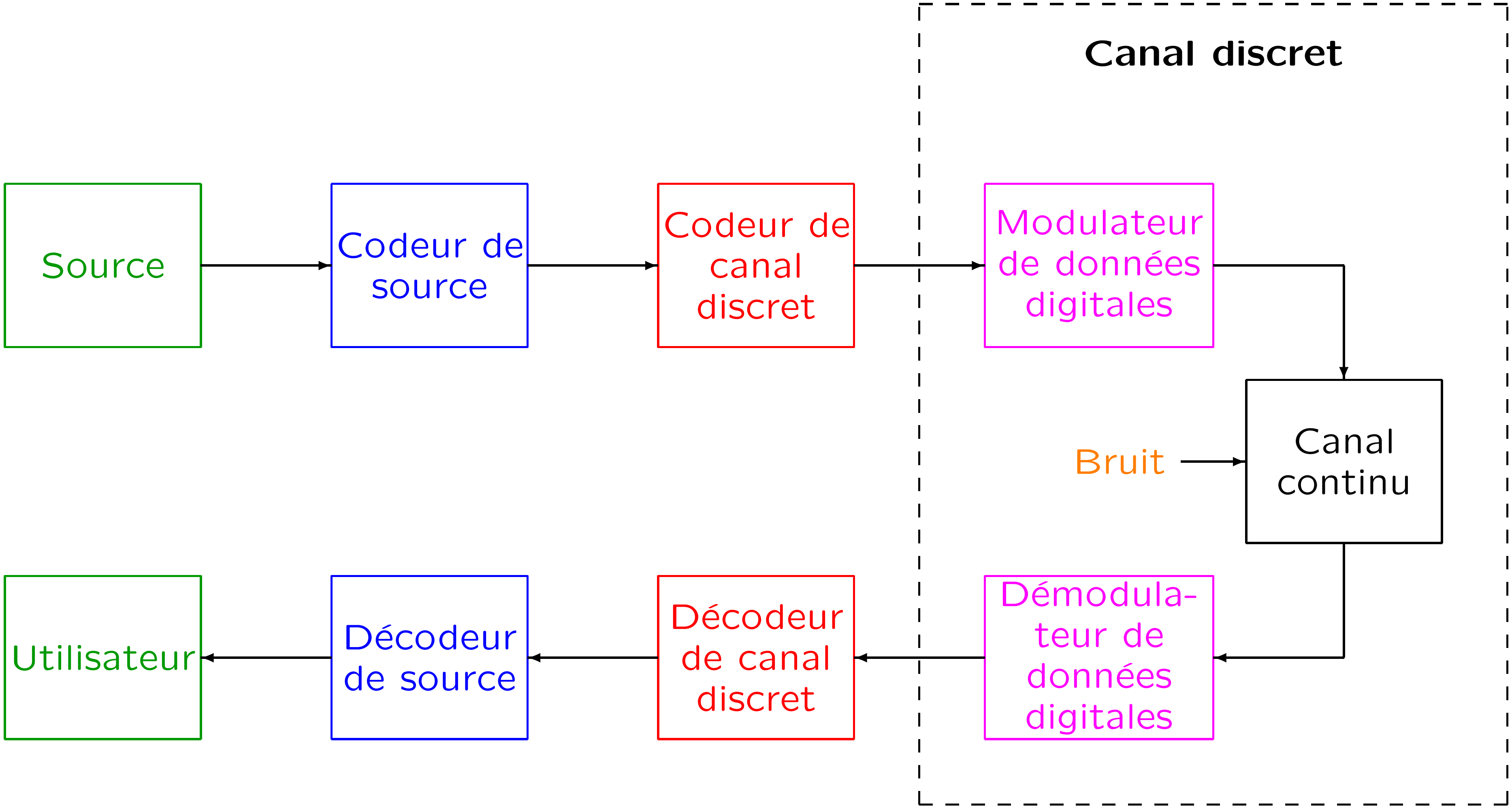
# Capacité du canal binaire symétrique




La capacité est le **taux de transmission maximal** du code à utiliser pour transmettre de l'information « dans de bonnes conditions ».

Par ex. $C(0.01) = 0.919$. Il y a donc moyen de faire (beaucoup) mieux que le code à répétition ! ! !

# Canal continu et canal discret



## Résultats importants du cours

**Premier théorème de Shannon** (Codage de source)

1. On peut coder toute source en utilisant un nombre de bits par lettre aussi proche que l'on veut de son entropie.
2. On ne peut pas faire mieux.

**Second théorème de Shannon** (Codage de canal)

1. On peut transmettre de l'information de façon fiable en utilisant un code correcteur d'erreur de taux de transmission inférieur à la capacité du canal utilisé.
2. On ne peut pas faire mieux.

*Ces résultats ne sont pas constructifs*