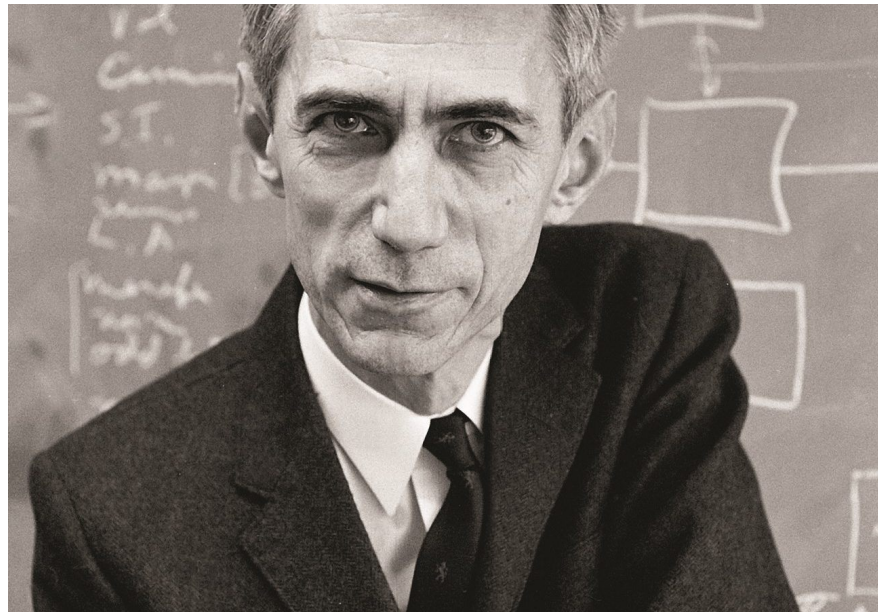


# Au sujet de *A mathematical Theory of communication,*

PAR CLAUDE ELWOOD SHANNON

*3 novembre 2022*



Quelques sources d'inspiration pour *A mathematical Theory of communication* (1948):

Quelques sources d'inspiration pour *A mathematical Theory of communication* (1948):

**Théorie des probabilités.** Notamment la formulation moderne de Andrey Kolmogorov (1935).

Quelques sources d'inspiration pour *A mathematical Theory of communication* (1948):

**Théorie des probabilités.** Notamment la formulation moderne de Andrey Kolmogorov (1935).

**Codes correcteurs d'erreurs.** Tels que présentés comme premiers exemples par Richard Hamming (1947).

Quelques sources d'inspiration pour *A mathematical Theory of communication* (1948):

**Théorie des probabilités.** Notamment la formulation moderne de Andrey Kolmogorov (1935).

**Codes correcteurs d'erreurs.** Tels que présentés comme premiers exemples par Richard Hamming (1947).

**Quantification statistique de l'information.** Selon les méthodes et points soulevés par John Tukey.

Quelques sources d'inspiration pour *A mathematical Theory of communication* (1948):

**Théorie des probabilités.** Notamment la formulation moderne de Andrey Kolmogorov (1935).

**Codes correcteurs d'erreurs.** Tels que présentés comme premiers exemples par Richard Hamming (1947).

**Quantification statistique de l'information.** Selon les méthodes et points soulevés par John Tukey.

**Notion(s) d'entropie.** L'entropie et ses propriétés telles qu'étudiées par Rudolf Clausius, William Gibbs, Ludwig Boltzmann et George Birkhoff ont servi d'inspiration à Shannon pour sa propre notion d'entropie.

Quelques sources d'inspiration pour *A mathematical Theory of communication* (1948):

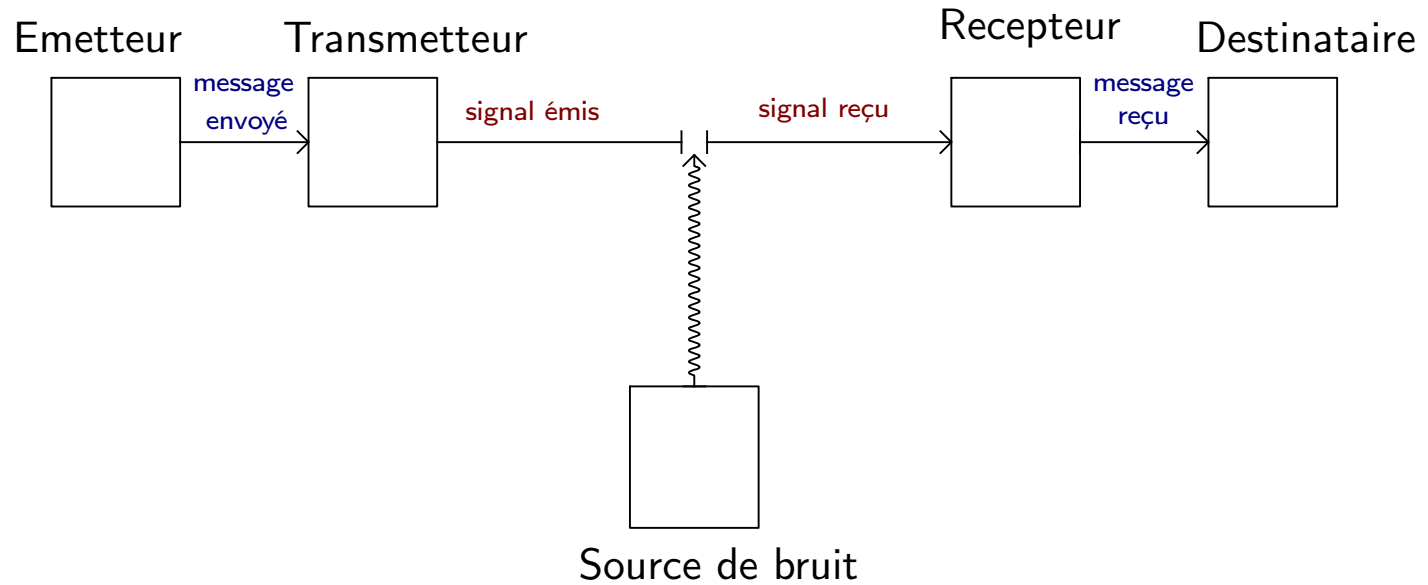
**Théorie des probabilités.** Notamment la formulation moderne de Andrey Kolmogorov (1935).

**Codes correcteurs d'erreurs.** Tels que présentés comme premiers exemples par Richard Hamming (1947).

**Quantification statistique de l'information.** Selon les méthodes et points soulevés par John Tukey.

**Notion(s) d'entropie.** L'entropie et ses propriétés telles qu'étudiées par Rudolf Clausius, William Gibbs, Ludwig Boltzmann et George Birkhoff ont servi d'inspiration à Shannon pour sa propre notion d'entropie.

On peut aussi citer Boole comme influence (posthume) concernant les centres d'intérêts, et von Neumann comme collègue qui semble avoir aidé à construire son approche pluridisciplinaire.











<b>Avant Shannon</b>	<b>Après Shannon</b>
bruit déterministe	bruit stochastique
chaîne de transmission comme objet d'étude	source comme objet d'étude
corriger les erreurs au niveau du signal	sélectionner de bonnes propriétés stochastiques
augmenter la fiabilité du réseau	étudier des algorithmes de compression

Nature du message	Exemples	Application	Signal
discret	0 1 1 1 0 1 0, chien rat lave	messages en langage courant; instructions logiques; télévision	signal discret (impul- sions électriques, sonores, lumineuses)
continu	fonction continue; éventuellement péri- odique; 	téléphonie et transmis- sion du son en général	signal continu: courant électrique variable
mixte	encodage de messages discrets <i>via</i> une modu- lation	contrôle basé sur des problèmes de syn- chronisation: contrôle aérien	signal continu (dans lequel des messages discrets sont encodés)

Nature du message	Exemples	Application	Signal
discret	0 1 1 1 0 1 0, chien rat lave	messages en langage courant; instructions logiques; télévision	signal discret (impulsions électriques, sonores, lumineuses)
continu	fonction continue; éventuellement périodique; 	téléphonie et transmission du son en général	signal continu: courant électrique variable
mixte	encodage de messages discrets <i>via</i> une modulation	contrôle basé sur des problèmes de synchronisation: contrôle aérien	signal continu (dans lequel des messages discrets sont encodés)

En pratique, les transformées de Fourier permettent de convertir un message continu en message discret et inversement:

$$f \rightarrow \hat{f} = \left( \omega \mapsto \int_0^{+\infty} f(t) e^{-it\omega} dt \right) \rightarrow (\hat{f}(\omega))_{\omega \in X} \rightarrow f(t) \approx \int_X \hat{f}(\omega) e^{-it\omega} d\omega.$$

On s'intéresse au support à travers lequel un signal discret est transmis. Le message est une combinaison finie quelconque des symboles  $S_1, \dots, S_n$ .

Pour  $t \geq 0$ , on note  $N(t)$  le nombre de messages qui peuvent être transmis en  $t$  secondes.

On s'intéresse au support à travers lequel un signal discret est transmis. Le message est une combinaison finie quelconque des symboles  $S_1, \dots, S_n$ .

Pour  $t \geq 0$ , on note  $N(t)$  le nombre de messages qui peuvent être transmis en  $t$  secondes.

Shannon définit la **capacité**  $C$  de la chaîne par

$$C := \lim_{t \rightarrow +\infty} \frac{N(t)}{t}.$$

Autrement dit, la moyenne sur le temps long du nombre de messages qui peuvent circuler sur la chaîne.

On s'intéresse au support à travers lequel un signal discret est transmis. Le message est une combinaison finie quelconque des symboles  $S_1, \dots, S_n$ .

Pour  $t \geq 0$ , on note  $N(t)$  le nombre de messages qui peuvent être transmis en  $t$  secondes.

Shannon définit la **capacité**  $C$  de la chaîne par

$$C := \lim_{t \rightarrow +\infty} \frac{N(t)}{t}.$$

Autrement dit, la moyenne sur le temps long du nombre de messages qui peuvent circuler sur la chaîne.

Remarque: lorsque  $t$  est grand,  $N(t)$  est arbitrairement proche de  $C(t) t$ .

$t_i$ : durée nécessaire à l'émission de  $S_i$  pour chaque  $1 \leq i \leq n$ .

Alors

$$N(t) = N(t - t_1) + N(t - t_2) + \cdots + N(t - t_n).$$

Résolution?



$t_i$ : durée nécessaire à l'émission de  $S_i$  pour chaque  $1 \leq i \leq n$ .

Alors

$$N = u_1(N) + u_2(N) + \cdots + u_n(N).$$

Où  $u_i(f) = t \mapsto f(t - t_i)$  est un « opérateur linéaire » et les  $u_i, 1 \leq i \leq n$  commutent 2 à 2.

$$u_i u_j = u_j u_i$$

$t_i$ : durée nécessaire à l'émission de  $S_i$  pour chaque  $1 \leq i \leq n$ .

Similaire à  $u_i = \partial^i$ , dérivée d'ordre  $i$ , dans le cas continu. L'algèbre linéaire générale donne

$$N(t) \approx \lambda^t,$$

où  $\lambda$  est la solution maximale de  $\lambda^{-t_1} + \lambda^{-t_2} + \dots + \lambda^{-t_n} = 1$ .

La contribution originale de Shannon est de s'intéresser (avec succès) à des propriétés statistiques des messages envoyés par la source d'information. Connaître ces propriétés permet de réduire le coût en information que requiert le décodage du signal reçu.

La contribution originale de Shannon est de s'intéresser (avec succès) à des propriétés statistiques des messages envoyés par la source d'information. Connaître ces propriétés permet de réduire le coût en information que requiert le décodage du signal reçu.

## Exemple: la source parle le français

*On sait si l'on reçoit le message OPPRI qu'il y a eu une ou plusieurs erreurs, et l'on peut chercher à déterminer où elles ont eu lieu et comment les corriger. Mais comment faire en pratique?*

- *En français, la lettre Z a une fréquence d'occurrence faible par rapport aux lettres E et A.*

La contribution originale de Shannon est de s'intéresser (avec succès) à des propriétés statistiques des messages envoyés par la source d'information. Connaître ces propriétés permet de réduire le coût en information que requiert le décodage du signal reçu.

## Exemple: la source parle le français

*On sait si l'on reçoit le message OPPRI qu'il y a eu une ou plusieurs erreurs, et l'on peut chercher à déterminer où elles ont eu lieu et comment les corriger. Mais comment faire en pratique?*

- *En français, la lettre Z a une fréquence d'occurrence faible par rapport aux lettres E et A.*
- *Ayant commencé un message avec la lettre L, on a plus de chance de le continuer avec un apostrophe ou une voyelle, qu'avec une consonne.*

La contribution originale de Shannon est de s'intéresser (avec succès) à des propriétés statistiques des messages envoyés par la source d'information. Connaître ces propriétés permet de réduire le coût en information que requiert le décodage du signal reçu.

## Exemple: la source parle le français

*On sait si l'on reçoit le message OPPRI qu'il y a eu une ou plusieurs erreurs, et l'on peut chercher à déterminer où elles ont eu lieu et comment les corriger. Mais comment faire en pratique?*

- *En français, la lettre Z a une fréquence d'occurrence faible par rapport aux lettres E et A.*
- *Ayant commencé un message avec la lettre L, on a plus de chance de le continuer avec un apostrophe ou une voyelle, qu'avec une consonne.*
- *Ayant reçu le mot GRAND, on a plus de chances de recevoir par la suite ARBRE que MANGER, MINUSCULE ou INSECTE.*

La contribution originale de Shannon est de s'intéresser (avec succès) à des propriétés statistiques des messages envoyés par la source d'information. Connaître ces propriétés permet de réduire le coût en information que requiert le décodage du signal reçu.

## Exemple: la source parle le français

*On sait si l'on reçoit le message OPPRI qu'il y a eu une ou plusieurs erreurs, et l'on peut chercher à déterminer où elles ont eu lieu et comment les corriger. Mais comment faire en pratique?*

- *En français, la lettre Z a une fréquence d'occurrence faible par rapport aux lettres E et A.*
- *Ayant commencé un message avec la lettre L, on a plus de chance de le continuer avec un apostrophe ou une voyelle, qu'avec une consonne.*
- *Ayant reçu le mot GRAND, on a plus de chances de recevoir par la suite ARBRE que MANGER, MINUSCULE ou INSECTE.*

En général, ayant reçu un début de signal  $\mathcal{S}_1$ , on a une certaine probabilité  $P(\mathcal{S}_1, \mathcal{S}_2)$  d'obtenir  $\mathcal{S}_2$  comme suite du signal, et cette probabilité n'est pas uniforme en  $\mathcal{S}_1, \mathcal{S}_2$ .

Shannon identifie la source à un processus stochastique: une chose pouvant être dans un ensemble fini d'états  $\mathcal{E}_1, \dots, \mathcal{E}_n$ , avec pour chaque  $1 \leq i, j \leq n$ , une probabilité  $p_{i,j} \in [0; 1]$  de passer de l'état  $\mathcal{E}_i$  à l'état  $\mathcal{E}_j$ .

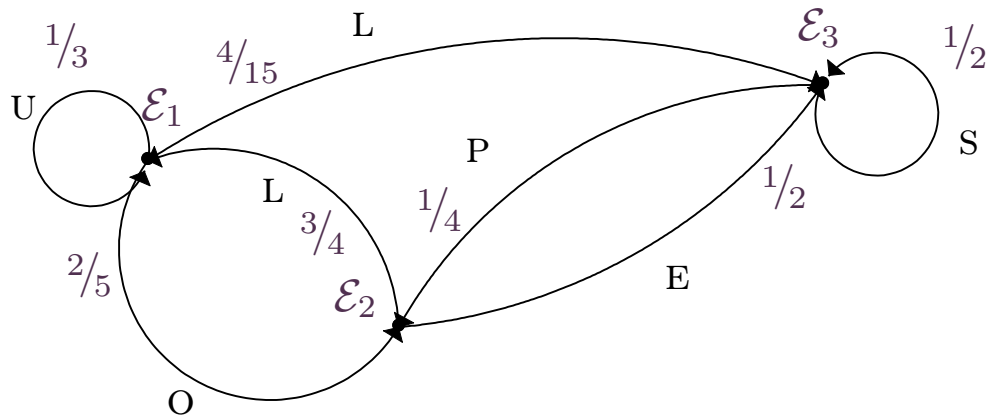


Shannon identifie la source à un processus stochastique: une chose pouvant être dans un ensemble fini d'états  $\mathcal{E}_1, \dots, \mathcal{E}_n$ , avec pour chaque  $1 \leq i, j \leq n$ , une probabilité  $p_{i,j} \in [0; 1]$  de passer de l'état  $\mathcal{E}_i$  à l'état  $\mathcal{E}_j$ .

A chaque transition  $\mathcal{E}_i \rightarrow \mathcal{E}_j$  un symbole  $S_{i,j}$  qui dépend uniquement de  $(i, j)$  est envoyé.

Shannon identifie la source à un processus stochastique: une chose pouvant être dans un ensemble fini d'états  $\mathcal{E}_1, \dots, \mathcal{E}_n$ , avec pour chaque  $1 \leq i, j \leq n$ , une probabilité  $p_{i,j} \in [0; 1]$  de passer de l'état  $\mathcal{E}_i$  à l'état  $\mathcal{E}_j$ .

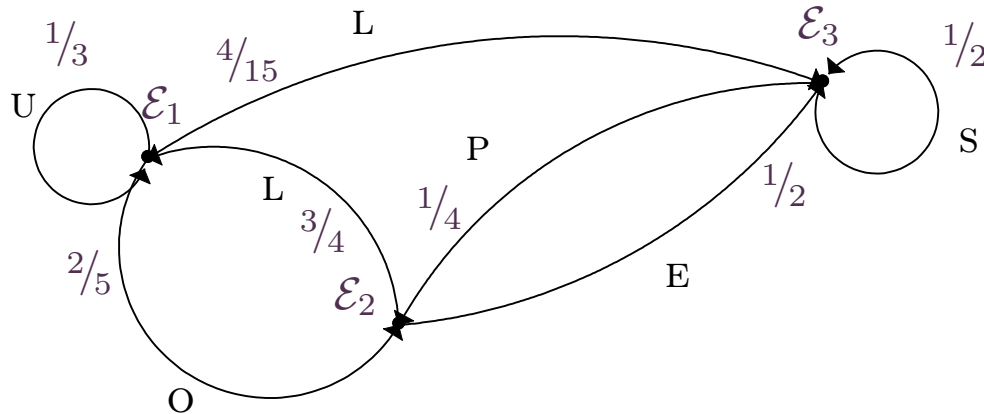
A chaque transition  $\mathcal{E}_i \rightarrow \mathcal{E}_j$  un symbole  $S_{i,j}$  qui dépend uniquement de  $(i, j)$  est envoyé.



Produit SEL avec une probabilité  $\frac{1}{2} \times \frac{1}{2} \times \frac{3}{4}$  à partir de l'état  $\mathcal{E}_3$ .

Shannon identifie la source à un processus stochastique: une chose pouvant être dans un ensemble fini d'états  $\mathcal{E}_1, \dots, \mathcal{E}_n$ , avec pour chaque  $1 \leq i, j \leq n$ , une probabilité  $p_{i,j} \in [0; 1]$  de passer de l'état  $\mathcal{E}_i$  à l'état  $\mathcal{E}_j$ .

A chaque transition  $\mathcal{E}_i \rightarrow \mathcal{E}_j$  un symbole  $S_{i,j}$  qui dépend uniquement de  $(i, j)$  est envoyé.



Produit SEL avec une probabilité  $\frac{1}{2} \times \frac{1}{2} \times \frac{3}{4}$  à partir de l'état  $\mathcal{E}_3$ .

Le système est **ergodique** lorsqu'il ne peut pas être coupé en deux systèmes stochastiques isolés (grosso modo). Cette notion vient de la théorie des systèmes dynamiques.

## Entropie d'une distribution statistique

Imaginons un ensemble  $\Omega$  d'évènements de probabilités  $p_1, \dots, p_n$  (avec  $p_1 + \dots + p_n = 1$ ).  
L'entropie de  $\Omega$  est le nombre

$$H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log(p_i).$$

Où  $\log(ab) = \log a + \log b$  pour tous  $a, b > 0$ , et  $\log(2^k) = k$  pour tout  $k \in \mathbb{N}$ .

ev<sub>1</sub>:  $N_1$  possibilités (ev<sub>1</sub>, ev<sub>2</sub>):  $N_1 \times N_2$  possibilités.

ev<sub>2</sub>:  $N_2$  possibilités



## Entropie d'une distribution statistique

Imaginons un ensemble  $\Omega$  d'évènements de probabilités  $p_1, \dots, p_n$  (avec  $p_1 + \dots + p_n = 1$ ).  
L'entropie de  $\Omega$  est le nombre

$$H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log(p_i).$$

Où  $\log(ab) = \log a + \log b$  pour tous  $a, b > 0$ , et  $\log(2^k) = k$  pour tout  $k \in \mathbb{N}$ .

Pourquoi cette formule?

ev<sub>1</sub>:  $N_1$  possibilités (ev<sub>1</sub>, ev<sub>2</sub>):  $N_1 \times N_2$  possibilités.

ev<sub>2</sub>:  $N_2$  possibilités



## Entropie d'une distribution statistique

Imaginons un ensemble  $\Omega$  d'évènements de probabilités  $p_1, \dots, p_n$  (avec  $p_1 + \dots + p_n = 1$ ).  
L'entropie de  $\Omega$  est le nombre

$$H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log(p_i).$$

Où  $\log(ab) = \log a + \log b$  pour tous  $a, b > 0$ , et  $\log(2^k) = k$  pour tout  $k \in \mathbb{N}$ .

Pourquoi cette formule?

$H(p_1, \dots, p_n)$  est maximale lorsque  $p_i = \frac{1}{n}$  pour tous les  $1 \leq i \leq n$ , lorsque la distribution est uniforme, et donc la moins informative possible.

ev<sub>1</sub>:  $N_1$  possibilités (ev<sub>1</sub>, ev<sub>2</sub>):  $N_1 \times N_2$  possibilités.

ev<sub>2</sub>:  $N_2$  possibilités





## Entropie d'une distribution statistique

Imaginons un ensemble  $\Omega$  d'évènements de probabilités  $p_1, \dots, p_n$  (avec  $p_1 + \dots + p_n = 1$ ).  
L'entropie de  $\Omega$  est le nombre

$$H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log(p_i).$$

Où  $\log(ab) = \log a + \log b$  pour tous  $a, b > 0$ , et  $\log(2^k) = k$  pour tout  $k \in \mathbb{N}$ .

Pourquoi cette formule?

$H(p_1, \dots, p_n)$  est maximale lorsque  $p_i = \frac{1}{n}$  pour tous les  $1 \leq i \leq n$ , lorsque la distribution est uniforme, et donc la moins informative possible.

Si l'on décompose l'évènement 1 en deux évènements de probabilités  $p_{1,1}$  et  $p_{1,2}$ , alors on obtient une entropie  $H(p_{1,1}, p_{1,2}, \dots, p_n) = H(p_1, \dots, p_n) + p_1 H(p_{1,1}, p_{1,2})$ :  $\equiv$  propriété d'extensivité de l'entropie en physique.

ev<sub>1</sub>:  $N_1$  possibilités (ev<sub>1</sub>, ev<sub>2</sub>):  $N_1 \times N_2$  possibilités.

ev<sub>2</sub>:  $N_2$  possibilités



On considère une source  $\mathcal{S} = \{\mathcal{E}_i, p_{i,j}, S_{i,j}\}$ . Pour chaque état  $\mathcal{E}_i$ , on a une probabilité  $P_i$  de tomber sur l'état  $\mathcal{E}_i$ .

Pour chaque symbole  $S \in \Sigma$ , on a une probabilité  $p_i(S) = \sum_{S_{i,j}=S} p_{i,j}$  d'obtenir le symbole  $S$  à partir de l'état  $\mathcal{E}_i$ . Donc une entropie  $H_i = H(p_i(S_0), \dots, p_i(S_m))$  correspondante.

On considère une source  $\mathcal{S} = \{\mathcal{E}_i, p_{i,j}, S_{i,j}\}$ . Pour chaque état  $\mathcal{E}_i$ , on a une probabilité  $P_i$  de tomber sur l'état  $\mathcal{E}_i$ .

Pour chaque symbole  $S \in \Sigma$ , on a une probabilité  $p_i(S) = \sum_{S_{i,j}=S} p_{i,j}$  d'obtenir le symbole  $S$  à partir de l'état  $\mathcal{E}_i$ . Donc une entropie  $H_i = H(p_i(S_0), \dots, p_i(S_m))$  correspondante.

## Formule pour l'entropie

*L'entropie de  $\mathcal{S}$  est définie par la formule*

$$H(\mathcal{S}) = - \sum_{i \leq n} P_i H_i.$$

*Cette fonction possède des propriétés similaires à l'entropie d'une distribution statistique.*

On considère une source  $\mathcal{S} = \{\mathcal{E}_i, p_{i,j}, S_{i,j}\}$ . Pour chaque état  $\mathcal{E}_i$ , on a une probabilité  $P_i$  de tomber sur l'état  $\mathcal{E}_i$ .

Pour chaque symbole  $S \in \Sigma$ , on a une probabilité  $p_i(S) = \sum_{S_{i,j}=S} p_{i,j}$  d'obtenir le symbole  $S$  à partir de l'état  $\mathcal{E}_i$ . Donc une entropie  $H_i = H(p_i(S_0), \dots, p_i(S_m))$  correspondante.

## Formule pour l'entropie temporelle

*L'entropie temporelle de  $\mathcal{S}$  est définie par la formule*

$$\tilde{H}(\mathcal{S}) = - \sum_{i \leq n} f_i H_i.$$

*où  $f_i$  est la fréquence temporelle moyenne de l'état  $\mathcal{E}_i$ .*

On considère une source  $\mathcal{S} = \{\mathcal{E}_i, p_{i,j}, S_{i,j}\}$ . Pour chaque état  $\mathcal{E}_i$ , on a une probabilité  $P_i$  de tomber sur l'état  $\mathcal{E}_i$ .

Pour chaque symbole  $S \in \Sigma$ , on a une probabilité  $p_i(S) = \sum_{S_{i,j}=S} p_{i,j}$  d'obtenir le symbole  $S$  à partir de l'état  $\mathcal{E}_i$ . Donc une entropie  $H_i = H(p_i(S_0), \dots, p_i(S_m))$  correspondante.

## Formule pour l'entropie temporelle

*L'entropie temporelle de  $\mathcal{S}$  est définie par la formule*

$$\tilde{H}(\mathcal{S}) = - \sum_{i \leq n} f_i H_i.$$

*où  $f_i$  est la fréquence temporelle moyenne de l'état  $\mathcal{E}_i$ .*

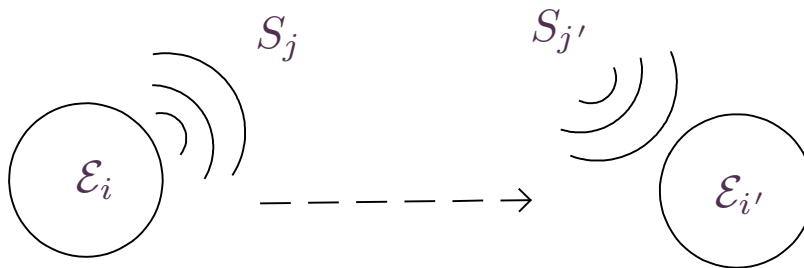
## Théorème (Shannon)

*Si une chaîne de capacité  $C$  transmet des signaux à partir de la source d'information  $\mathcal{S}$ , alors on a  $\tilde{H}(\mathcal{S}) \leq C$ . De plus, la chaîne étant donnée, il existe une source pour laquelle  $\tilde{H}(\mathcal{S}) = C$ .*

Supposons maintenant qu'un bruit perturbe la chaîne. Shannon s'intéresse en particulier à un bruit non déterministe. Ainsi, voyant la chaîne comme une source stochastique d'états  $\mathcal{E}_i$  et de symboles  $S_j$ , on a pour tout  $i, j, i', j'$  une probabilité

$$p_{i,j}(i', j')$$

que lorsque la chaîne est dans l'état  $\mathcal{E}_i$  et a envoyé le symbole  $S_j$ , on reçoit à cause du bruit le symbole  $S_{j'}$  avec une chaîne mise dans l'état  $\mathcal{E}_{i'}$ .

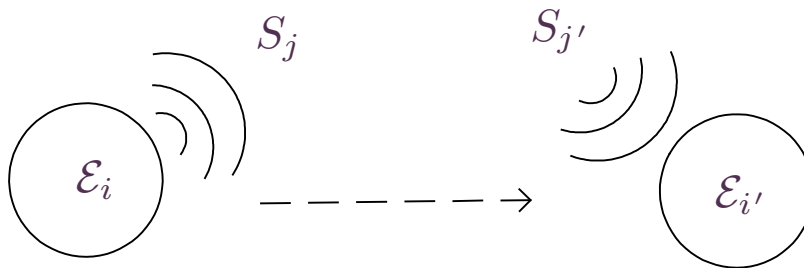




Supposons maintenant qu'un bruit perturbe la chaîne. Shannon s'intéresse en particulier à un bruit non déterministe. Ainsi, voyant la chaîne comme une source stochastique d'états  $\mathcal{E}_i$  et de symboles  $S_j$ , on a pour tout  $i, j, i', j'$  une probabilité

$$p_{i,j}(i', j')$$

que lorsque la chaîne est dans l'état  $\mathcal{E}_i$  et a envoyé le symbole  $S_j$ , on reçoit à cause du bruit le symbole  $S_{j'}$  avec une chaîne mise dans l'état  $\mathcal{E}_{i'}$ .



Comment définir la capacité de la chaîne dans ce cas?

## Probabilité conditionnelle

*Probabilité que  $B$  se réalise “sachant que”  $A$  se réalise:*

$$P_A(B) := \frac{P(A \text{ et } B)}{P(A)}.$$

## Probabilité conditionnelle

*Probabilité que  $B$  se réalise "sachant que"  $A$  se réalise:*

$$P_A(B) := \frac{P(A \text{ et } B)}{P(A)}.$$

On considère l'entropie  $H_{\mathcal{R}}(\mathcal{S})$  afin de mesurer l'information requise pour corriger un message après transmission. C'est-à-dire que dans la formule de  $H$ , on remplace les probabilités par les probabilités conditionnelles.

## Probabilité conditionnelle

Probabilité que  $B$  se réalise “sachant que”  $A$  se réalise:

$$P_A(B) := \frac{P(A \text{ et } B)}{P(A)}.$$

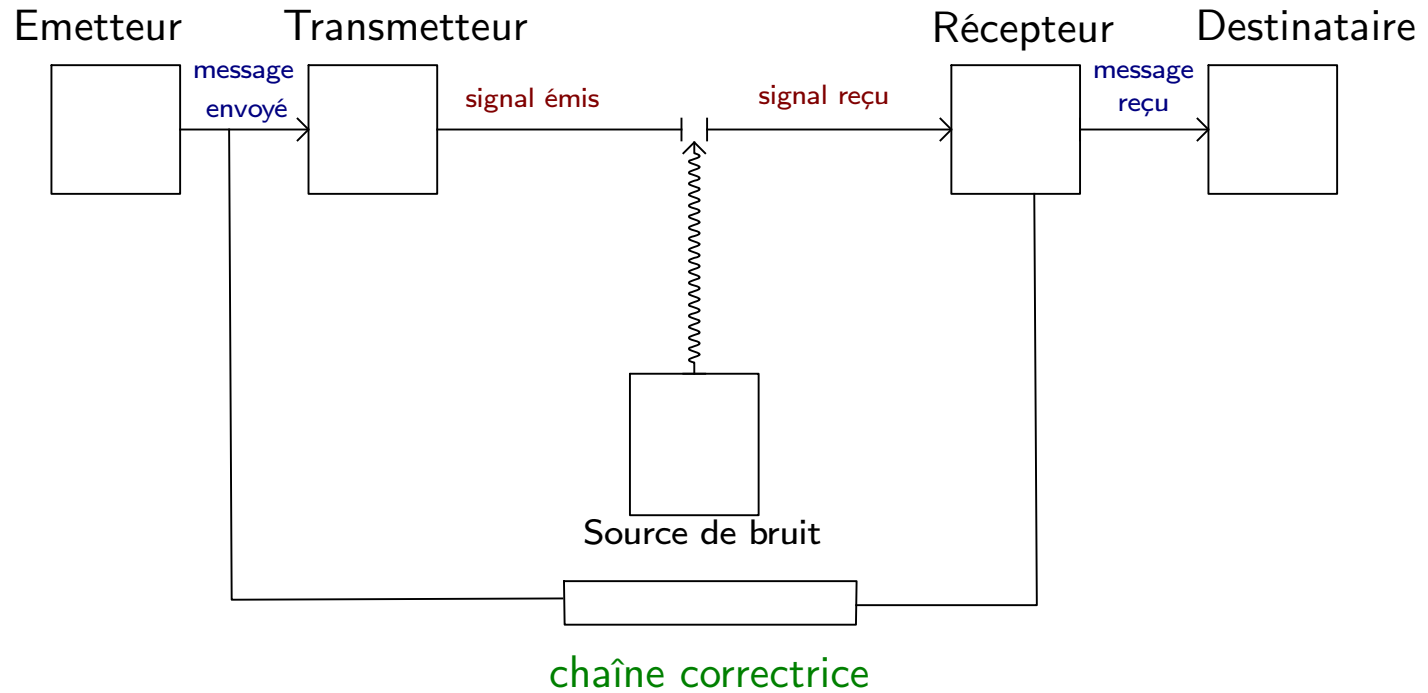
On considère l'entropie  $H_{\mathcal{R}}(\mathcal{S})$  afin de mesurer l'information requise pour corriger un message après transmission. C'est-à-dire que dans la formule de  $H$ , on remplace les probabilités par les probabilités conditionnelles.

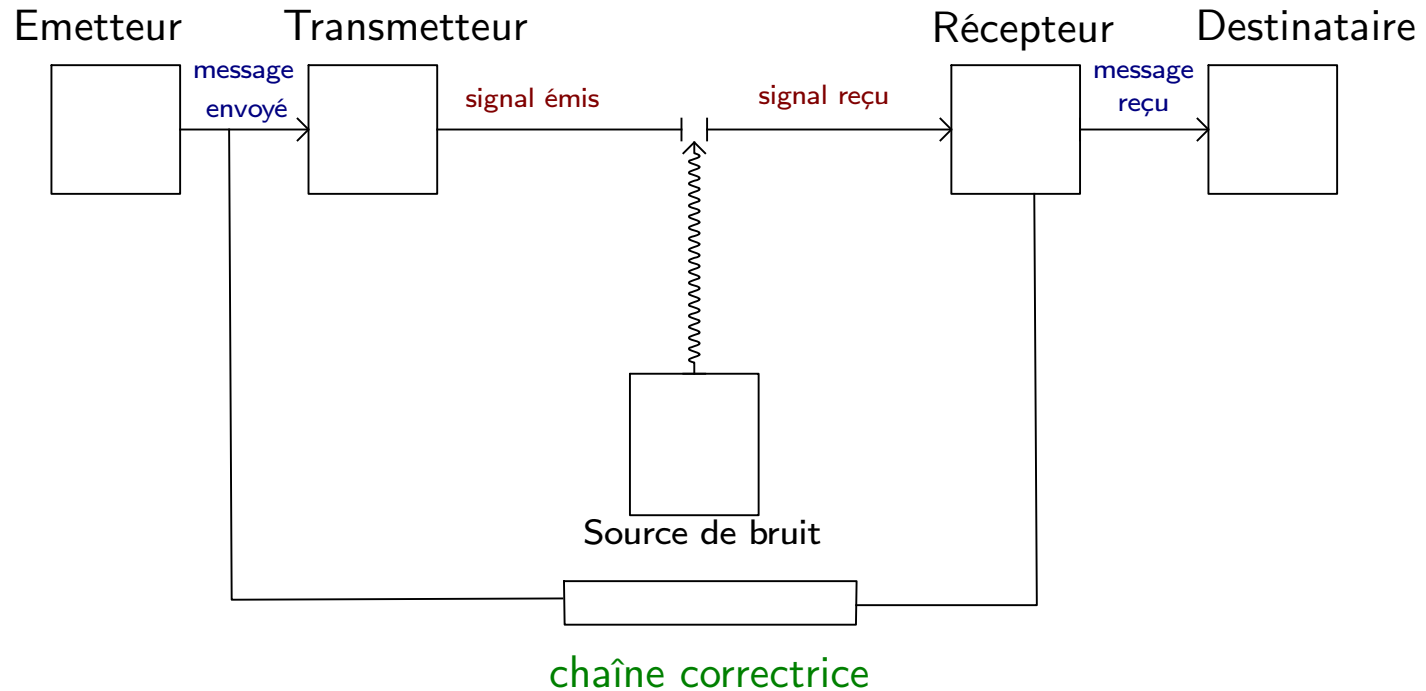
On peut montrer que  $H_{\mathcal{R}}(\mathcal{S}) \leq H(\mathcal{S})$ , avec égalité si et seulement si les messages reçus sont indépendants de ceux envoyés (par exemple à cause d'un bruit trop fort).

On définit alors la capacité d'une chaîne avec bruit en considérant toutes les sources possibles  $\mathcal{S}$  pour la chaîne, et en posant

$$C := \max_{\mathcal{S}} (\tilde{H}(\mathcal{S}) - \tilde{H}_{\mathcal{R}}(\mathcal{S})),$$

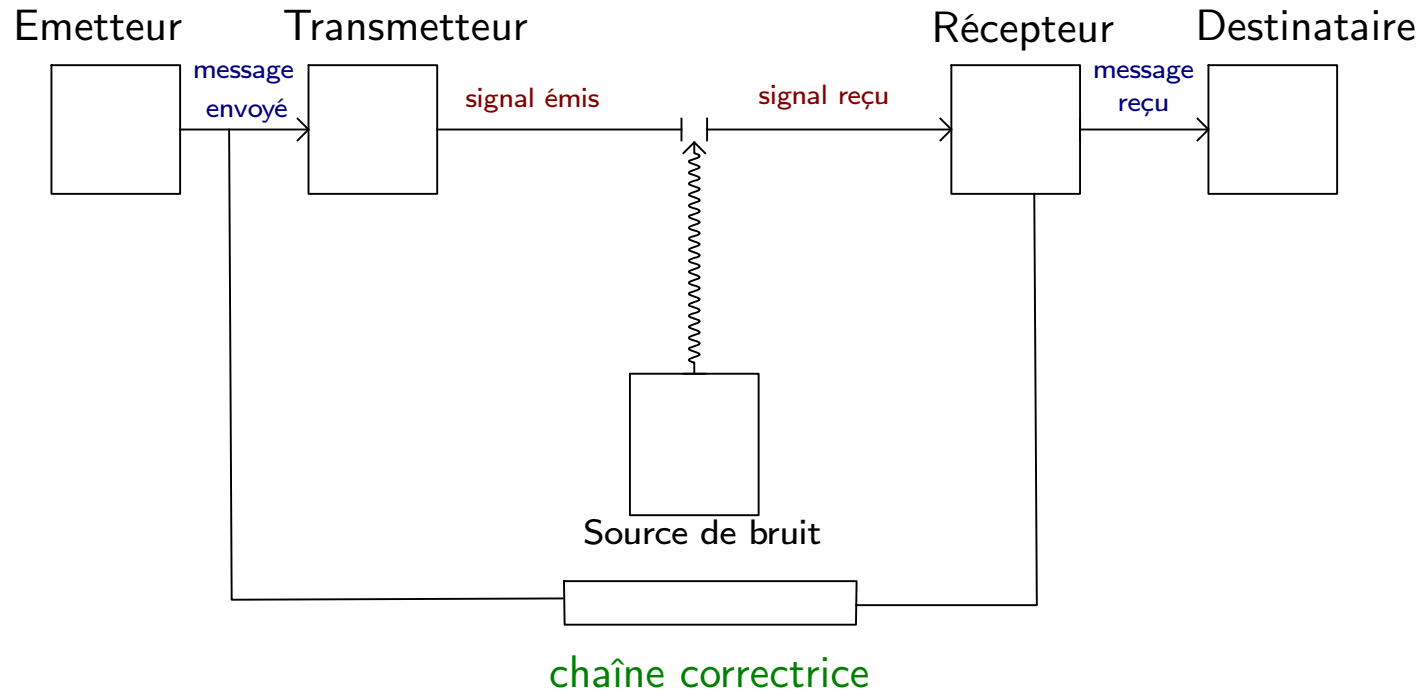
ce qui est cohérent avec le cas sans bruit (où  $\tilde{H}_{\mathcal{R}}(\mathcal{S}) = 0$ ).





## Théorèmes 10 et 11

- Une chaîne de correction de capacité  $\tilde{H}_{\mathcal{R}}(\mathcal{S})$  peut corriger avec un taux d'échec arbitrairement bas les erreurs d'une source  $\mathcal{S}$ . Cette capacité est minimale.



## Théorèmes 10 et 11

- Une chaîne de correction de capacité  $\tilde{H}_{\mathcal{R}}(\mathcal{S})$  peut corriger avec un taux d'échec arbitrairement bas les erreurs d'une source  $\mathcal{S}$ . Cette capacité est minimale.
- Une chaîne de capacité  $C$  peut transmettre les messages d'une source  $\mathcal{S}$  avec un taux d'erreur arbitrairement bas si et seulement si  $\tilde{H}(\mathcal{S}) \leq C$ .

Dans le cas d'un signal continu, les idées de Shannon s'appliquent, mais il n'avait pas à l'époque les outils théoriques pour pouvoir prouver les résultats dans le même ordre de généralité.

Il faut retenir que l'on peut définir l'entropie d'une source d'information, et qu'elle possède des propriétés similaires à l'entropie dans le cas discret.

Cependant Shannon utilise le cas discret pour approximer le cas continu, et en l'état ses résultats suggèrent que le modèle est plus adapté à la transmission de messages discrets.



Le travail de Shannon a deux effets:

- introduire le caractère aléatoire (processus stochastiques) maîtrisé (ergodicité) comme outil pour donner des limites théoriques optimales
- abstraire la notion de source d'information et donc permettre une certaine liberté et hauteur de vue

Le travail de Shannon a deux effets:

- introduire le caractère aléatoire (processus stochastiques) maîtrisé (ergodicité) comme outil pour donner des limites théoriques optimales
- abstraire la notion de source d'information et donc permettre une certaine liberté et hauteur de vue

On peut considérer le travail de Shannon comme le début à proprement parler de la théorie de l'information.

Le travail de Shannon a deux effets:

- introduire le caractère aléatoire (processus stochastiques) maîtrisé (ergodicité) comme outil pour donner des limites théoriques optimales
- abstraire la notion de source d'information et donc permettre une certaine liberté et hauteur de vue

On peut considérer le travail de Shannon comme le début à proprement parler de la théorie de l'information.

Les théorèmes 10 et 11 ouvrent la possibilité d'étudier des notions générales de compression d'information, l'utilité de la redondance d'information, le lien entre bande passante et capacité à corriger des erreurs...