Cours 1. Introduction. Source d'information.

A. Désilles

29 mars 2010

Résumé

Résumé

Résumé

Résumé

Résumé

Résumé

Plan du cours

Plan du cours

Plan du cours

Plan du cours

Plan du cours

Plan du cours

Objectifs du cours

- Introduire les concepts principaux de la théorie de l'information
- Présenter les principaux domaines d'application
- Initier aux techniques de base de codage

Objectifs du cours

- Introduire les concepts principaux de la théorie de l'information
- Présenter les principaux domaines d'application
- Initier aux techniques de base de codage

Objectifs du cours

- Introduire les concepts principaux de la théorie de l'information
- Présenter les principaux domaines d'application
- Initier aux techniques de base de codage

Un peu d'histoire

- **1838** Le premier télégraphe électrique, conçu par S. Morse fonctionne.
- Chaque lettre de l'alphabet est représentée par une séquence de
  - points (courent électrique de courte durée)
  - traits (courent de longue durée)
  - espaces (absence de courent)
- Fait curieux : la lettre "E" est représentée de la façon la plus courte possible : un point.

Un peu d'histoire

- **1838** Le premier télégraphe électrique, conçu par S. Morse fonctionne.
- Chaque lettre de l'alphabet est représentée par une séquence de
  - **points** (courent électrique de courte durée)
  - **traits** (courent de longue durée)
  - **espaces** (absence de courent)
- Fait curieux : la lettre "E" est représentée de la façon la plus courte possible : un point.

Un peu d'histoire

- **1838** Le premier télégraphe électrique, conçu par S. Morse fonctionne.
- Chaque lettre de l'alphabet est représentée par une séquence de
  - **points** (courent électrique de courte durée)
  - **traits** (courent de longue durée)
  - **espaces** (absence de courent)
- Fait curieux : la lettre "E" est représentée de la façon la plus courte possible : un point.

Un peu d'histoire

- **1838** Le premier télégraphe électrique, conçu par S. Morse fonctionne.
- Chaque lettre de l'alphabet est représentée par une séquence de
  - **points** (courent électrique de courte durée)
  - **traits** (courent de longue durée)
  - **espaces** (absence de courent)
- **Fait curieux :** la lettre **"E"** est représentée de la façon la plus courte possible : un point.

Un peu d'histoire

- **1838** Le premier télégraphe électrique, conçu par S. Morse fonctionne.
- Chaque lettre de l'alphabet est représentée par une séquence de
  - **points** (courent électrique de courte durée)
  - **traits** (courent de longue durée)
  - **espaces** (absence de courent)
- **Fait curieux :** la lettre **"E"** est représentée de la façon la plus courte possible : un point.

Un peu d'histoire

- **1838** Le premier télégraphe électrique, conçu par S. Morse fonctionne.
- Chaque lettre de l'alphabet est représentée par une séquence de
  - **points** (courent électrique de courte durée)
  - **traits** (courent de longue durée)
  - **espaces** (absence de courent)
- **Fait curieux :** la lettre **"E"** est représentée de la façon la plus courte possible : un point.

Un peu d'histoire

- **1843** Construction de ligne télégraphique entre Washington et Baltimore
- Les lignes sont enterrées
- Des problèmes électriques causent des erreurs de transmission

Un peu d'histoire

- **1843** Construction de ligne télégraphique entre Washington et Baltimore
- Les lignes sont enterrées
- Des problèmes électriques causent des erreurs de transmission

Un peu d'histoire

- **1843** Construction de ligne télégraphique entre Washington et Baltimore
- Les lignes sont enterrées
- Des problèmes électriques causent des erreurs de transmission

Incertitude dans un processus de communication

- On ne connaît pas à l'avance le message émis. $\Rightarrow$ 1ère source d'incertitude.
- On ne connaît pas les erreurs qui ont été commises lors de la transmission. $\Rightarrow$ 2ème source d'incertitude.

Question

Comment quantifier l'incertitude ?

Incertitude dans un processus de communication

- On ne connaît pas à l'avance le message émis. $\Rightarrow$ 1ère source d'incertitude.
- On ne connaît pas les erreurs qui ont été commises lors de la transmission. $\Rightarrow$ 2ème source d'incertitude.

Question

Comment quantifier l'incertitude ?

Incertitude dans un processus de communication

- On ne connaît pas à l'avance le message émis. $\Rightarrow$ 1ère source d'incertitude.
- On ne connaît pas les erreurs qui ont été commises lors de la transmission. $\Rightarrow$ 2ème source d'incertitude.

Question

Comment quantifier l'incertitude ?

Incertitude dans un processus de communication

- On ne connaît pas à l'avance le message émis. $\Rightarrow$ 1ère source d'incertitude.
- On ne connaît pas les erreurs qui ont été commises lors de la transmission. $\Rightarrow$ 2ème source d'incertitude.

Question

Comment quantifier l'incertitude ?

Information = Incertitude ?

- En théorie de la communication le terme d'information est associé à l'incertitude qui accompagne l'émission et la transmission d'un message.
- Ainsi, mesurer l'information signifie mesurer l'incertitude.

Problématique

La théorie de l'information fournit un formalisme rigoureux qui permet de modéliser, quantifier et analyser la notion de l'information.

Information = Incertitude ?

- En théorie de la communication le terme d'information est associé à l'incertitude qui accompagne l'émission et la transmission d'un message.
- Ainsi, mesurer l'information signifie mesurer l'incertitude.

Problématique

La théorie de l'information fournit un formalisme rigoureux qui permet de modéliser, quantifier et analyser la notion de l'information.

Information = Incertitude ?

- En théorie de la communication le terme d'information est associé à l'incertitude qui accompagne l'émission et la transmission d'un message.
- Ainsi, mesurer l'information signifie mesurer l'incertitude.

Problématique

La théorie de l'information fournit un formalisme rigoureux qui permet de modéliser, quantifier et analyser la notion de l'information.

Information = Incertitude ?

- En théorie de la communication le terme d'information est associé à l'incertitude qui accompagne l'émission et la transmission d'un message.
- Ainsi, mesurer l'information signifie mesurer l'incertitude.

Problématique

La théorie de l'information fournit un formalisme rigoureux qui permet de modéliser, quantifier et analyser la notion de l'information.

Théorie de l'information et applications

- Le codage est une mise en application de la théorie de l'information. La théorie de codage développe les techniques pratiques de stockage et de transmission d'informations fiables.
- Compression des données. La théorie de l'information permet d'établir les limites théoriques de compression sans pertes.
- Cryptographie. Les concepts de la théorie de l'information permettent de définir et d'évaluer la sécurité des systèmes cryptographiques

Théorie de l'information et applications

- Le codage est une mise en application de la théorie de l'information. La théorie de codage développe les techniques pratiques de stockage et de transmission d'informations fiables.
- Compression des données. La théorie de l'information permet d'établir les limites théoriques de compression sans pertes.
- Cryptographie. Les concepts de la théorie de l'information permettent de définir et d'évaluer la sécurité des systèmes cryptographiques

Théorie de l'information et applications

- Le codage est une mise en application de la théorie de l'information. La théorie de codage développe les techniques pratiques de stockage et de transmission d'informations fiables.
- Compression des données. La théorie de l'information permet d'établir les limites théoriques de compression sans pertes.
- Cryptographie. Les concepts de la théorie de l'information permettent de définir et d'évaluer la sécurité des systèmes cryptographiques

Modèle de communication

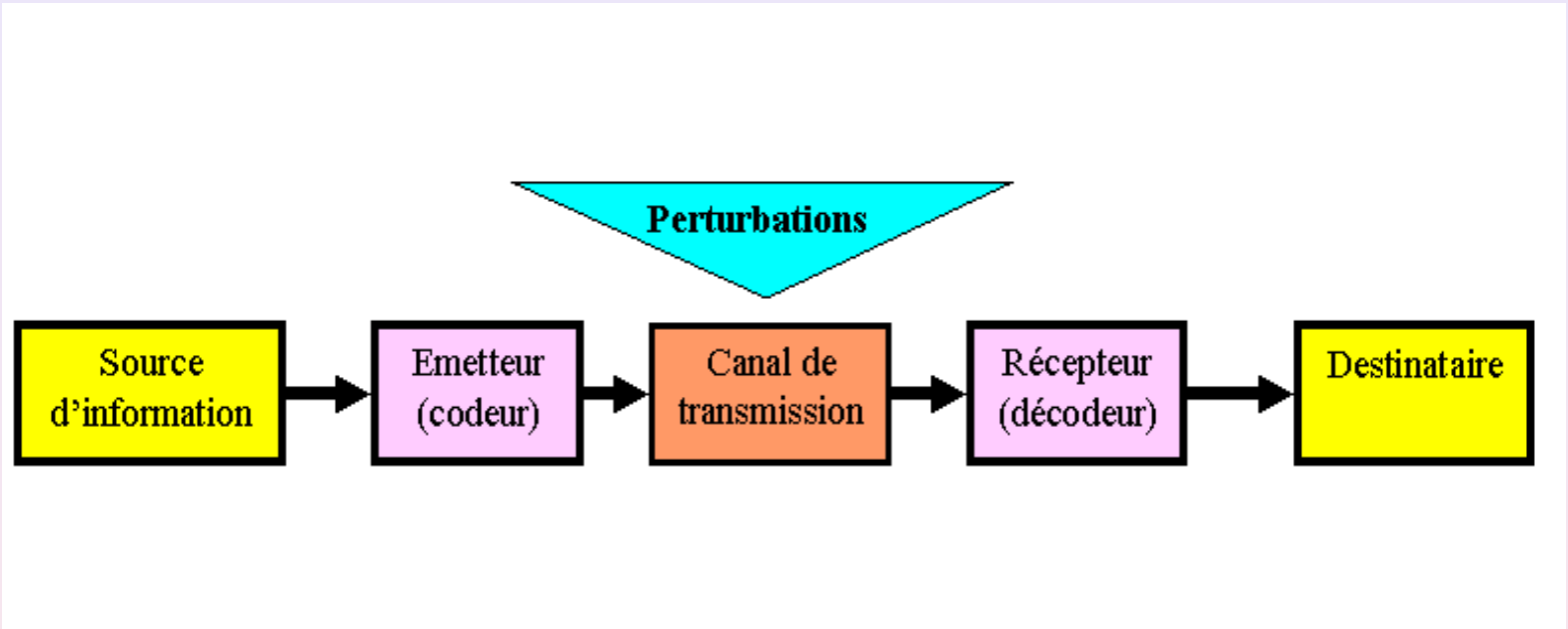


Figure: Paradigme de Shannon

Paradigme de Shannon

- la **source d'information** choisit un message M parmi un certain nombre de messages possibles ;
- l'**émetteur** transforme le message en signal S compatible physiquement avec le mode de transmission choisi.
- le signal S est alors soumis à l'entrée d'un **canal de transmission** ;
- lors de la transmission des perturbations peuvent transformer le signal envoyé ; on parle alors **de bruit de canal** ;
- à la sortie du canal le signal $\tilde{S}$ est soumis au décodeur qui le transforme en message $\tilde{M}$ lisible par le **destinataire**.

Paradigme de Shannon

- la **source d'information** choisit un message M parmi un certain nombre de messages possibles ;
- l'**émetteur** transforme le message en signal S compatible physiquement avec le mode de transmission choisi.
- le signal S est alors soumis à l'entrée d'un **canal de transmission** ;
- lors de la transmission des perturbations peuvent transformer le signal envoyé ; on parle alors **de bruit de canal** ;
- à la sortie du canal le signal $\tilde{S}$ est soumis au décodeur qui le transforme en message $\tilde{M}$ lisible par le **destinataire**.

Paradigme de Shannon

- la **source d'information** choisit un message M parmi un certain nombre de messages possibles ;
- l'**émetteur** transforme le message en signal S compatible physiquement avec le mode de transmission choisi.
- le signal S est alors soumis à l'entrée d'un **canal de transmission** ;
- lors de la transmission des perturbations peuvent transformer le signal envoyé ; on parle alors **de bruit de canal** ;
- à la sortie du canal le signal $\tilde{S}$ est soumis au décodeur qui le transforme en message $\tilde{M}$ lisible par le **destinataire**.

Paradigme de Shannon

- la **source d'information** choisit un message M parmi un certain nombre de messages possibles ;
- l'**émetteur** transforme le message en signal S compatible physiquement avec le mode de transmission choisi.
- le signal S est alors soumis à l'entrée d'un **canal de transmission** ;
- lors de la transmission des perturbations peuvent transformer le signal envoyé ; on parle alors **de bruit de canal** ;
- à la sortie du canal le signal $\tilde{S}$ est soumis au décodeur qui le transforme en message $\tilde{M}$ lisible par le **destinataire**.

Paradigme de Shannon

- la **source d'information** choisit un message M parmi un certain nombre de messages possibles ;
- l'**émetteur** transforme le message en signal S compatible physiquement avec le mode de transmission choisi.
- le signal S est alors soumis à l'entrée d'un **canal de transmission** ;
- lors de la transmission des perturbations peuvent transformer le signal envoyé ; on parle alors **de bruit de canal** ;
- à la sortie du canal le signal $\tilde{S}$ est soumis au décodeur qui le transforme en message $\tilde{M}$ lisible par le **destinataire**.

Caractère aléatoire d'une source

- Du point de vue du destinataire le message émis n'est pas connu à l'avance.
- Il peut donc être considéré comme étant aléatoire.
- On peut définir une variable aléatoire X associée à l'observation d'un symbole émis par la source.
- L'alphabet $\Omega_X = \{x_1, \ldots, x_n\}$ représente alors le support (l'ensemble de valeurs possibles de X)
- L'ensemble de probabilités $P_X = \{p_i,\ i = 1, \ldots, n\}$ d'émission des symboles définit la distribution de probabilités de X.

Caractère aléatoire d'une source

- Du point de vue du destinataire le message émis n'est pas connu à l'avance.
- Il peut donc être considéré comme étant aléatoire.
- On peut définir une variable aléatoire X associée à l'observation d'un symbole émis par la source.
- L'alphabet $\Omega_X = \{x_1, \ldots, x_n\}$ représente alors le support (l'ensemble de valeurs possibles de X)
- L'ensemble de probabilités $P_X = \{p_i,\ i = 1, \ldots, n\}$ d'émission des symboles définit la distribution de probabilités de X.

Caractère aléatoire d'une source

- Du point de vue du destinataire le message émis n'est pas connu à l'avance.
- Il peut donc être considéré comme étant aléatoire.
- On peut définir une variable aléatoire X associée à l'observation d'un symbole émis par la source.
- L'alphabet $\Omega_X = \{x_1, \ldots, x_n\}$ représente alors le support (l'ensemble de valeurs possibles de X)
- L'ensemble de probabilités $P_X = \{p_i,\ i = 1, \ldots, n\}$ d'émission des symboles définit la distribution de probabilités de X.

Caractère aléatoire d'une source

- Du point de vue du destinataire le message émis n'est pas connu à l'avance.
- Il peut donc être considéré comme étant aléatoire.
- On peut définir une variable aléatoire X associée à l'observation d'un symbole émis par la source.
- L'alphabet $\Omega_X = \{x_1, \ldots, x_n\}$ représente alors le support (l'ensemble de valeurs possibles de X)
- L'ensemble de probabilités $P_X = \{p_i,\ i = 1, \ldots, n\}$ d'émission des symboles définit la distribution de probabilités de X.

Caractère aléatoire d'une source

- Du point de vue du destinataire le message émis n'est pas connu à l'avance.
- Il peut donc être considéré comme étant aléatoire.
- On peut définir une variable aléatoire X associée à l'observation d'un symbole émis par la source.
- L'alphabet $\Omega_X = \{x_1, \ldots, x_n\}$ représente alors le support (l'ensemble de valeurs possibles de X)
- L'ensemble de probabilités $P_X = \{p_i,\ i = 1, \ldots, n\}$ d'émission des symboles définit la distribution de probabilités de X.

Caractère aléatoire d'un canal de transmission

- Selon le modèle de communication de Shannon, le message émis par la source est soumis à l'entrée d'un canal de transmission ;
- Les perturbations présentes dans le canal modifient le signal initial ;
- Ces perturbations ne sont pas connues à l'avance par le destinataire ;
- Ainsi la transmission d'un message a également un caractère aléatoire.

Caractère aléatoire d'un canal de transmission

- Selon le modèle de communication de Shannon, le message émis par la source est soumis à l'entrée d'un canal de transmission ;
- Les perturbations présentes dans le canal modifient le signal initial ;
- Ces perturbations ne sont pas connues à l'avance par le destinataire ;
- Ainsi la transmission d'un message a également un caractère aléatoire.

Caractère aléatoire d'un canal de transmission

- Selon le modèle de communication de Shannon, le message émis par la source est soumis à l'entrée d'un canal de transmission ;
- Les perturbations présentes dans le canal modifient le signal initial ;
- Ces perturbations ne sont pas connues à l'avance par le destinataire ;
- Ainsi la transmission d'un message a également un caractère aléatoire.

Caractère aléatoire d'un canal de transmission

- Selon le modèle de communication de Shannon, le message émis par la source est soumis à l'entrée d'un canal de transmission ;
- Les perturbations présentes dans le canal modifient le signal initial ;
- Ces perturbations ne sont pas connues à l'avance par le destinataire ;
- Ainsi la transmission d'un message a également un caractère aléatoire.

Modèle d'une source d'information

Une source d'information X est décrite par un couple (Ω_X, P_X) où Ω_X est un alphabet fini et P_X est une distribution de probabilités sur Ω_X.

Exemple

Soit une source binaire, produisant des symboles 0 et 1 avec les probabilités respectives p et $q = 1 - p$. Nous avons ici l'alphabet constitué de deux symboles $\Omega = \{0, 1\}$ et $P[X = 0] = p, \quad P[X = 1] = q$.

Modèle d'une source d'information

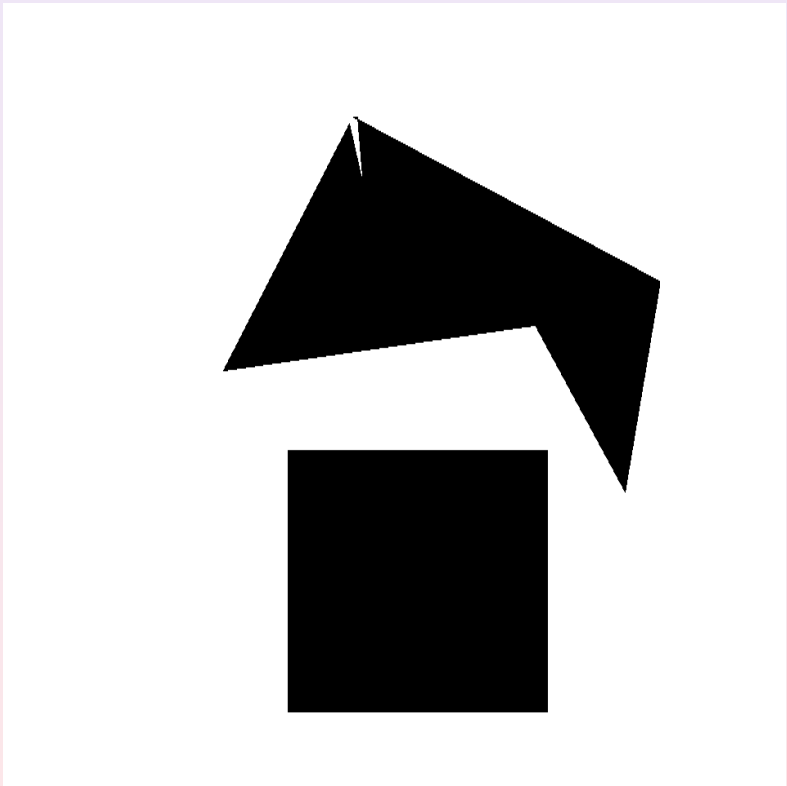
Une source d'information X est décrite par un couple (Ω_X, P_X) où Ω_X est un alphabet fini et P_X est une distribution de probabilités sur Ω_X.

Exemple

Soit une source binaire, produisant des symboles 0 et 1 avec les probabilités respectives p et $q = 1 - p$. Nous avons ici l'alphabet constitué de deux symboles $\Omega = \{0, 1\}$ et $P[X = 0] = p, \quad P[X = 1] = q$.

Un autre exemple

Cette image est un message d'une source binaire : 0 pour le noir et 1 pour le blanc



Un texte est aussi une source d'informations

Un texte français est un message d'une source utilisant l'alphabet latin. La distribution de probabilité de cette source est définie par les fréquences d'apparition des lettres, établies empiriquement.

Lettre	e	s	a	i	t	n	r	u
Probabilité (%)	14.715	7.948	7.636	7.529	7.244	7.095	6.553	6.311

Information propre.

- Que cherche-t-on à mesurer exactement ?
- La difficulté que l'on a à prévoir un événement.
- On joue à "pile ou face" ? Comment quantifiez vous votre difficulté à deviner le résultat ?
- Et si on lance un dé à 6 faces ? Est ce plus difficile à deviner ?
- Et si on vous donnait une indication ? Par exemple, que le résultat est pair.
- Combien d'information cette indication vous a apporté ?
- Intuitivement, on pourrait dire : "plus l'événement est rare, plus il est difficile à prévoir"

Information propre.

- Que cherche-t-on à mesurer exactement ?
- La difficulté que l'on a à prévoir un événement.
- On joue à "pile ou face" ? Comment quantifiez vous votre difficulté à deviner le résultat ?
- Et si on lance un dé à 6 faces ? Est ce plus difficile à deviner ?
- Et si on vous donnait une indication ? Par exemple, que **le résultat est pair**.
- Combien d'information cette indication vous a apporté ?
- Intuitivement, on pourrait dire : "plus l'événement est rare, plus il est difficile à prévoir"

Information propre.

- Que cherche-t-on à mesurer exactement ?
- La difficulté que l'on a à prévoir un événement.
- On joue à "pile ou face" ? Comment quantifiez vous votre difficulté à deviner le résultat ?
- Et si on lance un dé à 6 faces ? Est ce plus difficile à deviner ?
- Et si on vous donnait une indication ? Par exemple, que **le résultat est pair**.
- **Combien** d'information cette indication vous a apporté ?
- Intuitivement, on pourrait dire : "plus l'événement est rare, plus il est difficile à prévoir"

Information propre.

- Que cherche-t-on à mesurer exactement ?
- La difficulté que l'on a à prévoir un événement.
- On joue à "pile ou face" ? Comment quantifiez vous votre difficulté à deviner le résultat ?
- Et si on lance un dé à 6 faces ? Est ce plus difficile à deviner ?
- Et si on vous donnait une indication ? Par exemple, que le résultat est pair.
- Combien d'information cette indication vous a apporté ?
- Intuitivement, on pourrait dire : "plus l'événement est rare, plus il est difficile à prévoir"

Information propre.

- Que cherche-t-on à mesurer exactement ?
- La difficulté que l'on a à prévoir un événement.
- On joue à "pile ou face" ? Comment quantifiez vous votre difficulté à deviner le résultat ?
- Et si on lance un dé à 6 faces ? Est ce plus difficile à deviner ?
- Et si on vous donnait une indication ? Par exemple, que **le résultat est pair**.
- **Combien** d'information cette indication vous a apporté ?
- Intuitivement, on pourrait dire : "plus l'événement est rare, plus il est difficile à prévoir"

Information propre.

- Que cherche-t-on à mesurer exactement ?
- La difficulté que l'on a à prévoir un événement.
- On joue à "pile ou face" ? Comment quantifiez vous votre difficulté à deviner le résultat ?
- Et si on lance un dé à 6 faces ? Est ce plus difficile à deviner ?
- Et si on vous donnait une indication ? Par exemple, que **le résultat est pair**.
- **Combien** d'information cette indication vous a apporté ?
- Intuitivement, on pourrait dire : "plus l'événement est rare, plus il est difficile à prévoir"

Information propre.

- Que cherche-t-on à mesurer exactement ?
- La difficulté que l'on a à prévoir un événement.
- On joue à "pile ou face" ? Comment quantifiez vous votre difficulté à deviner le résultat ?
- Et si on lance un dé à 6 faces ? Est ce plus difficile à deviner ?
- Et si on vous donnait une indication ? Par exemple, que **le résultat est pair**.
- **Combien** d'information cette indication vous a apporté ?
- Intuitivement, on pourrait dire : "plus l'événement est rare, plus il est difficile à prévoir"

Un premier exemple : probabilités et information

- Reprenons l'exemple du dé.
- Imaginons que le résultat est $X = 2$. Soit A l'événement associé. $P(A) = \frac{1}{6}$.
- Est ce difficile à deviner ? Nous avons une chance de $\frac{1}{6}$ d'avoir la réponse correcte.
- Imaginons que l'on nous dit que le résultat du lancé est un nombre pair. Soit B l'événement associé. On a $P(B) = \frac{1}{2}$.
- Cette information, augmente notre chance de réussite.
- En effet, nous avons maintenant, la probabilité de $P(A|B) = \frac{1}{3}$ de deviner le résultat.

Un premier exemple : probabilités et information

- Reprenons l'exemple du dé.
- Imaginons que le résultat est $X = 2$. Soit A l'événement associé. $P(A) = \frac{1}{6}$.
- Est ce difficile à deviner ? Nous avons une chance de $\frac{1}{6}$ d'avoir la réponse correcte.
- Imaginons que l'on nous dit que le résultat du lancé est un nombre pair. Soit B l'événement associé. On a $P(B) = \frac{1}{2}$.
- Cette information, augmente notre chance de réussite.
- En effet, nous avons maintenant, la probabilité de $P(A|B) = \frac{1}{3}$ de deviner le résultat.

Un premier exemple : probabilités et information

- Reprenons l'exemple du dé.
- Imaginons que le résultat est $X = 2$. Soit A l'événement associé. $P(A) = \frac{1}{6}$.
- Est ce difficile à deviner ? Nous avons une chance de $\frac{1}{6}$ d'avoir la réponse correcte.
- Imaginons que l'on nous dit que le résultat du lancé est un nombre pair. Soit B l'événement associé. On a $P(B) = \frac{1}{2}$.
- Cette information, augmente notre chance de réussite.
- En effet, nous avons maintenant, la probabilité de $P(A|B) = \frac{1}{3}$ de deviner le résultat.

Un premier exemple : probabilités et information

- Reprenons l'exemple du dé.
- Imaginons que le résultat est $X = 2$. Soit A l'événement associé. $P(A) = \frac{1}{6}$.
- Est ce difficile à deviner ? Nous avons une chance de $\frac{1}{6}$ d'avoir la réponse correcte.
- Imaginons que l'on nous dit que le résultat du lancé est un nombre pair. Soit B l'événement associé. On a $P(B) = \frac{1}{2}$.
- Cette information, augmente notre chance de réussite.
- En effet, nous avons maintenant, la probabilité de $P(A|B) = \frac{1}{3}$ de deviner le résultat.

Un premier exemple : probabilités et information

- Reprenons l'exemple du dé.
- Imaginons que le résultat est $X = 2$. Soit A l'événement associé. $P(A) = \frac{1}{6}$.
- Est ce difficile à deviner ? Nous avons une chance de $\frac{1}{6}$ d'avoir la réponse correcte.
- Imaginons que l'on nous dit que le résultat du lancé est un nombre pair. Soit B l'événement associé. On a $P(B) = \frac{1}{2}$.
- Cette information, augmente notre chance de réussite.
- En effet, nous avons maintenant, la probabilité de $P(A|B) = \frac{1}{3}$ de deviner le résultat.

Un premier exemple : probabilités et information

- Reprenons l'exemple du dé.
- Imaginons que le résultat est $X = 2$. Soit A l'événement associé. $P(A) = \frac{1}{6}$.
- Est ce difficile à deviner ? Nous avons une chance de $\frac{1}{6}$ d'avoir la réponse correcte.
- Imaginons que l'on nous dit que le résultat du lancé est un nombre pair. Soit B l'événement associé. On a $P(B) = \frac{1}{2}$.
- Cette information, augmente notre chance de réussite.
- En effet, nous avons maintenant, la probabilité de $P(A|B) = \frac{1}{3}$ de deviner le résultat.

Premier exemple : conclusion

L'observation d'un événement aléatoire A nous apporte une quantité d'information $h(A)$ qui est

- non négative : $h(A) \geq 0$
- nulle si l'événement est certain : $P(A) = 1 \Rightarrow h(A) = 0$
- d'autant plus grande que l'événement est rare ou incertain : $P(A_1) < P(A_2) \Rightarrow h(A_1) > h(A_2)$
- additive pour les événements indépendants : $h(A \cap B) = h(A) + h(B)$

Premier exemple : conclusion

L'observation d'un événement aléatoire A nous apporte une quantité d'information $h(A)$ qui est

- non négative : $h(A) \geq 0$
- nulle si l'événement est certain : $P(A) = 1 \Rightarrow h(A) = 0$
- d'autant plus grande que l'événement est rare ou incertain : $P(A_1) < P(A_2) \Rightarrow h(A_1) > h(A_2)$
- additive pour les événements indépendants : $h(A \cap B) = h(A) + h(B)$

Premier exemple : conclusion

L'observation d'un événement aléatoire A nous apporte une quantité d'information $h(A)$ qui est

- non négative : $h(A) \geq 0$
- nulle si l'événement est certain : $P(A) = 1 \Rightarrow h(A) = 0$
- d'autant plus grande que l'événement est rare ou incertain : $P(A_1) < P(A_2) \Rightarrow h(A_1) > h(A_2)$
- additive pour les événements indépendants : $h(A \cap B) = h(A) + h(B)$

Premier exemple : conclusion

L'observation d'un événement aléatoire A nous apporte une quantité d'information $h(A)$ qui est

- non négative : $h(A) \geq 0$
- nulle si l'événement est certain : $P(A) = 1 \Rightarrow h(A) = 0$
- d'autant plus grande que l'événement est rare ou incertain : $P(A_1) < P(A_2) \Rightarrow h(A_1) > h(A_2)$
- additive pour les événements indépendants : $h(A \cap B) = h(A) + h(B)$

Information propre. Définition

Information propre

Soient $\Omega = \{\omega_1, \ldots, \omega_m\}$ un alphabet discret et X la variable aléatoire associée. Pour tout événement $A \subset \Omega$ la quantité d'information propre de A est définie par

$$h(A) = -\log_2(P(A)).$$

Entropie d'une source. Idée.

Soit une source X d'alphabet $\Omega_X = \{x_1, \ldots, x_m\}$ et de distribution de probabilité $P_X = \{p_1, \ldots, p_n\}$.
L'information associé à l'observation de chaque symbole x_i est $-\log_2(p_i)$

Quantité d'information moyenne

L'entropie d'une telle source représente la quantité moyenne d'information propre associée à l'observation de chacun des symboles possibles.

Entropie d'une source. Idée.

Soit une source X d'alphabet $\Omega_X = \{x_1, \ldots, x_m\}$ et de distribution de probabilité $P_X = \{p_1, \ldots, p_n\}$.
L'information associé à l'observation de chaque symbole x_i est $-\log_2(p_i)$

Quantité d'information moyenne

L'entropie d'une telle source représente la quantité moyenne d'information propre associée à l'observation de chacun des symboles possibles.

Entropie d'une source. Idée.

Soit une source X d'alphabet $\Omega_X = \{x_1, \ldots, x_m\}$ et de distribution de probabilité $P_X = \{p_1, \ldots, p_n\}$.
L'information associé à l'observation de chaque symbole x_i est $-\log_2(p_i)$

Quantité d'information moyenne

L'entropie d'une telle source représente la quantité moyenne d'information propre associée à l'observation de chacun des symboles possibles.

Entropie d'une source. Définition.

Entropie d'une source

Soient $\Omega_X = \{x_1, \ldots, x_m\}$ l'alphabet fini d'une source et X la variable aléatoire associée t.q. $P[\omega_i] = p_i, \quad i = 1, \ldots, m$. On appelle **entropie** ou encore **quantité moyenne d'information** de la source la quantité

$$H(X) = H(p_1, p_2, \ldots, p_n) = E[h(x)] = -\sum_{i=1}^{m} p_i \log_2(p_i)$$

L'unité de mesure de cette quantité est le "bit par symbole".

Exemple. Entropie d'une source binaire

Soit une source émettant des symboles 0 avec la probabilité p et 1 avec la probabilité $q = 1 - p$.

$$H_2(p) = -p \log_2(p) - (1-p) \log_2(1-p).$$

Lorsque les deux symboles sont équiprobables ($p = 1/2$) :

$$H_2(1/2) = -1/2 \log_2(1/2) - 1/2 \log_2(1/2) = 1.$$

Nous pouvons interpréter ce résultat comme suit : *lorsque les symboles d'une source binaire sont équiprobables, il faut un bit par symbole en moyenne.*

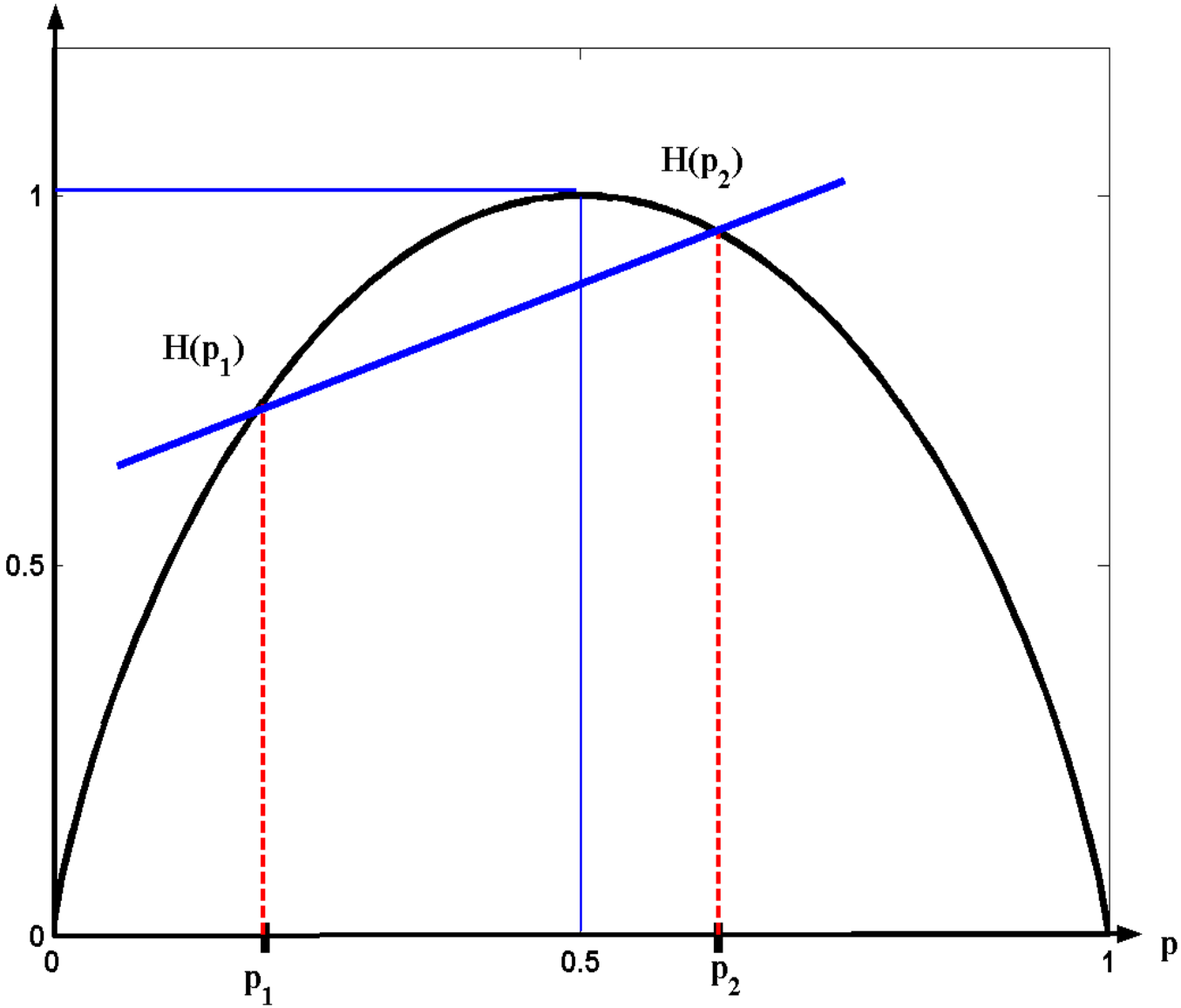


Figure: Fonction d'entropie d'une source binaire

Propriétés de l'entropie d'une source binaire

Soit $H_2(p),\ p \in [0,1]$ l'entropie d'une source binaire.

1. $H_2(p)$ est une fonction continue sur $]0,1[$ telle que

$$\lim_{p \to 0+} H_2(p) = 0 = \lim_{p \to 1-} H_2(p)$$

2. $H_2(p)$ est positive sur $H_2(p),\ p \in [0,1]$
3. $H_2(p)$ est symétrique par rapport à $p_0 = 0.5$ et atteint son maximum en p_0 t.q. $H_2(0.5) = 1$.

Propriétés de l'entropie d'une source binaire

Soit $H_2(p),\ p \in [0,1]$ l'entropie d'une source binaire.

1. $H_2(p)$ est une fonction continue sur $]0,1[$ telle que

$$\lim_{p \to 0+} H_2(p) = 0 = \lim_{p \to 1-} H_2(p)$$

2. $H_2(p)$ est positive sur $H_2(p),\ p \in [0,1]$
3. $H_2(p)$ est symétrique par rapport à $p_0 = 0.5$ et atteint son maximum en p_0 t.q. $H_2(0.5) = 1$.

Propriétés de l'entropie d'une source binaire

Soit $H_2(p),\ p \in [0,1]$ l'entropie d'une source binaire.

1. $H_2(p)$ est une fonction continue sur $]0,1[$ telle que

$$\lim_{p \to 0+} H_2(p) = 0 = \lim_{p \to 1-} H_2(p)$$

2. $H_2(p)$ est positive sur $H_2(p),\ p \in [0,1]$
3. $H_2(p)$ est symétrique par rapport à $p_0 = 0.5$ et atteint son maximum en p_0 t.q. $H_2(0.5) = 1$.

Exemple

Soit une source S définie par

X	a	b	c	d	e
P(X)	0.3	0.2	0.2	0.15	0.15

L'entropie de cette source, calculée selon la définition est

Exemple

Soit une source S définie par

X	a	b	c	d	e
P(X)	0.3	0.2	0.2	0.15	0.15

L'entropie de cette source, calculée selon la définition est

Exemple

Soit une source S définie par

X	a	b	c	d	e
P(X)	0.3	0.2	0.2	0.15	0.15

L'entropie de cette source, calculée selon la définition est

$$H(X) = -\sum_{i=1}^{5} p_i \log_2(p_i) =$$

Exemple

Soit une source S définie par

X	a	b	c	d	e
P(X)	0.3	0.2	0.2	0.15	0.15

L'entropie de cette source, calculée selon la définition est

$$H(X) = -\sum_{i=1}^{5} p_i \log_2(p_i) = -(0.3 \log_2(0.3)$$

Exemple

Soit une source S définie par

X	a	b	c	d	e
P(X)	0.3	0.2	0.2	0.15	0.15

L'entropie de cette source, calculée selon la définition est

$$H(X) = -\sum_{i=1}^{5} p_i \log_2(p_i) = -(0.3 \log_2(0.3) - 2 \cdot 0.2 \log_2(0.2)$$

Exemple

Soit une source S définie par

X	a	b	c	d	e
P(X)	0.3	0.2	0.2	0.15	0.15

L'entropie de cette source, calculée selon la définition est

$$H(X) = -\sum_{i=1}^{5} p_i \log_2(p_i) = -(0.3 \log_2(0.3) - 2 \cdot 0.2 \log_2(0.2) - 2 \cdot 0.15 \log_2(0.15$$

Exemple

Soit une source S définie par

X	a	b	c	d	e
P(X)	0.3	0.2	0.2	0.15	0.15

L'entropie de cette source, calculée selon la définition est

$$H(X) = -\sum_{i=1}^{5} p_i \log_2(p_i) = -(0.3 \log_2(0.3) - 2 \cdot 0.2 \log_2(0.2) - 2 \cdot 0.15 \log_2(0.15$$

$$H(X) \simeq 2.27.$$

Entropie d'une source. Propriétés

Soit X une source d'alphabet $\Omega_X = \{x_i\}_{i=1}^n$et de de distribution de probabilité P donnée $P[X = x_i] = p_i, \quad i = 1, \ldots, n$. Notons $H(p_1, \ldots, p_n)$ sa fonction d'entropie. Alors

- **Positivité**

$$H(p_1, p_2, \ldots, p_n) \geq 0.$$

L'égalité a lieu uniquement si l'une des probabilités p_i est égale à 1 et les autres sont nulles.

- **Propriété de maximum**

$$H(p_1, p_2, \ldots, p_n) \leq log(n)$$

et l'égalité a lieu si et seulement si $\forall i = 1, \ldots, n, \quad p_i = \frac{1}{n}$.

- A retenir. L'entropie est maximale lorsque la distribution des probabilités est uniforme : tous les symboles sont équiprobables.

Entropie d'une source. Propriétés

Soit X une source d'alphabet $\Omega_X = \{x_i\}_{i=1}^n$et de de distribution de probabilité P donnée $P[X = x_i] = p_i, \quad i = 1, \ldots, n$. Notons $H(p_1, \ldots, p_n)$ sa fonction d'entropie. Alors

- **Positivité**

$$H(p_1, p_2, \ldots, p_n) \geq 0.$$

  L'égalité a lieu uniquement si l'une des probabilités p_i est égale à 1 et les autres sont nulles.

- **Propriété de maximum**

$$H(p_1, p_2, \ldots, p_n) \leq log(n)$$

  et l'égalité a lieu si et seulement si $\forall i = 1, \ldots, n, \quad p_i = \frac{1}{n}$.

- A retenir. L'entropie est maximale lorsque la distribution des probabilités est uniforme : tous les symboles sont équiprobables.

Entropie d'une source. Propriétés

Soit X une source d'alphabet $\Omega_X = \{x_i\}_{i=1}^n$et de de distribution de probabilité P donnée $P[X = x_i] = p_i, \quad i = 1, \ldots, n$. Notons $H(p_1, \ldots, p_n)$ sa fonction d'entropie. Alors

- **Positivité**

$$H(p_1, p_2, \ldots, p_n) \geq 0.$$

  L'égalité a lieu uniquement si l'une des probabilités p_i est égale à 1 et les autres sont nulles.

- **Propriété de maximum**

$$H(p_1, p_2, \ldots, p_n) \leq log(n)$$

  et l'égalité a lieu si et seulement si $\forall i = 1, \ldots, n, \quad p_i = \frac{1}{n}$.

- A retenir. L'entropie est maximale lorsque la distribution des probabilités est uniforme : tous les symboles sont équiprobables.

Entropie conjointe

Définition

Soient $X = \{x_i\}_{i=1}^n$ et $Y = \{y_j\}_{j=1}^m$ deux variables aléatoires discrètes définies sur un même univers.
Soit $P(i,j) = P[X = x_i \text{ et } Y = y_j]$ leur distribution conjointe.
Alors **l'entropie conjointe de X et Y** est définie par

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{n} P(i,j)\log(P(i,j)).$$

Il s'agit de l'entropie de la variable aléatoire $Z = (X,Y)$ dont les valeurs sont les couples (x_i, y_j).

Entropie conjointe

Définition

Soient $X = \{x_i\}_{i=1}^n$ et $Y = \{y_j\}_{j=1}^m$ deux variables aléatoires discrètes définies sur un même univers.
Soit $P(i,j) = P[X = x_i \text{ et } Y = y_j]$ leur distribution conjointe.
Alors **l'entropie conjointe de X et Y** est définie par

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{n} P(i,j)\log(P(i,j)).$$

Il s'agit de l'entropie de la variable aléatoire $Z = (X,Y)$ dont les valeurs sont les couples (x_i, y_j).

Entropie conditionnelle moyenne

Définition

Soient $X = \{x_i\}_{i=1}^n$ et $Y = \{y_j\}_{j=1}^m$ deux variables aléatoires discrètes définies sur un même univers.

Soit $P(i,j) = P[X = x_i \text{et } Y = y_j]$ leur distribution conjointe.

Posons $P(i|j) = P[X = x_i | Y = y_j] = \dfrac{P(i,j)}{P[Y = y_j]}$.

Alors l'**entropie conditionnelle moyenne** de X sachant Y est définie par

$$H(X|Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P(i,j)\log(P(i|j)).$$

Cryptographie. Vocabulaire.

- La **cryptographie** est une science qui étudie les méthodes permettant de transmettre des messages de façon confidentielle.
- **message clair** : un ensemble de données (texte, image,...) que l'on souhaite transmettre.
- Le **chiffrement** est un procédé permettant de transformer le message clair de telle sorte qu'il soit incompréhensible par qui que ce soit d'autre que l'auteur du message et le destinataire.
- Le **chiffré** est le résultat du chiffrement.
- Le **déchiffrement** est le procédé permettant de retrouver le message clair à partir du chiffré.
- La **clé** de chiffrement est un paramètre qui est utilisé dans le procédé de chiffrement ou déchiffrement.

Cryptographie. Vocabulaire.

- La **cryptographie** est une science qui étudie les méthodes permettant de transmettre des messages de façon confidentielle.
- **message clair** : un ensemble de données (texte, image,...) que l'on souhaite transmettre.
- Le **chiffrement** est un procédé permettant de transformer le message clair de telle sorte qu'il soit incompréhensible par qui que ce soit d'autre que l'auteur du message et le destinataire.
- Le **chiffré** est le résultat du chiffrement.
- Le **déchiffrement** est le procédé permettant de retrouver le message clair à partir du chiffré.
- La clé de chiffrement est un paramètre qui est utilisé dans le procédé de chiffrement ou déchiffrement.

Cryptographie. Vocabulaire.

- La **cryptographie** est une science qui étudie les méthodes permettant de transmettre des messages de façon confidentielle.
- **message clair** : un ensemble de données (texte, image,...) que l'on souhaite transmettre.
- Le **chiffrement** est un procédé permettant de transformer le message clair de telle sorte qu'il soit incompréhensible par qui que ce soit d'autre que l'auteur du message et le destinataire.
- Le **chiffré** est le résultat du chiffrement.
- Le **déchiffrement** est le procédé permettant de retrouver le message clair à partir du chiffré.
- La **clé** de chiffrement est un paramètre qui est utilisé dans le procédé de chiffrement ou déchiffrement.

Cryptographie. Vocabulaire.

- La **cryptographie** est une science qui étudie les méthodes permettant de transmettre des messages de façon confidentielle.
- **message clair** : un ensemble de données (texte, image,...) que l'on souhaite transmettre.
- Le **chiffrement** est un procédé permettant de transformer le message clair de telle sorte qu'il soit incompréhensible par qui que ce soit d'autre que l'auteur du message et le destinataire.
- Le **chiffré** est le résultat du chiffrement.
- Le **déchiffrement** est le procédé permettant de retrouver le message clair à partir du chiffré.
- La **clé** de chiffrement est un paramètre qui est utilisé dans le procédé de chiffrement ou déchiffrement.

Cryptographie. Vocabulaire.

- La **cryptographie** est une science qui étudie les méthodes permettant de transmettre des messages de façon confidentielle.
- **message clair** : un ensemble de données (texte, image,...) que l'on souhaite transmettre.
- Le **chiffrement** est un procédé permettant de transformer le message clair de telle sorte qu'il soit incompréhensible par qui que ce soit d'autre que l'auteur du message et le destinataire.
- Le **chiffré** est le résultat du chiffrement.
- Le **déchiffrement** est le procédé permettant de retrouver le message clair à partir du chiffré.
- La **clé** de chiffrement est un paramètre qui est utilisé dans le procédé de chiffrement ou déchiffrement.

Cryptographie. Vocabulaire.

- La **cryptographie** est une science qui étudie les méthodes permettant de transmettre des messages de façon confidentielle.
- **message clair** : un ensemble de données (texte, image,...) que l'on souhaite transmettre.
- Le **chiffrement** est un procédé permettant de transformer le message clair de telle sorte qu'il soit incompréhensible par qui que ce soit d'autre que l'auteur du message et le destinataire.
- Le **chiffré** est le résultat du chiffrement.
- Le **déchiffrement** est le procédé permettant de retrouver le message clair à partir du chiffré.
- La **clé** de chiffrement est un paramètre qui est utilisé dans le procédé de chiffrement ou déchiffrement.

Cryptosystème

Définition

Un **cryptosystème** est un quintuplé $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ formé de

- un ensemble de clés $\mathcal{K}$
- un ensemble de messages clairs possibles , $\mathcal{M}$,
- un ensemble de messages chiffrés possibles $\mathcal{C}$
- un algorithme de chiffrage, représenté par une fonction $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$,
- et un procédé de déchiffrage , représenté par une fonction $D : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$.

On suppose que pour tout $m \in \mathcal{M}$ il existe une paide de clés de chiffrement et de déchiffrement telles que la relation

$$D(k_D, E(k_E, m)) = m$$

est assurée.

Cryptosystème

Définition

Un **cryptosystème** est un quintuplé $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ formé de

- un ensemble de clés $\mathcal{K}$
- un ensemble de messages clairs possibles , $\mathcal{M}$,
- un ensemble de messages chiffrés possibles $\mathcal{C}$
- un algorithme de chiffrage, représenté par une fonction $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$,
- et un procédé de déchiffrage , représenté par une fonction $D : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$.

On suppose que pour tout $m \in \mathcal{M}$ il existe une paide de clés de chiffrement et de déchiffrement telles que la relation

$$D(k_D, E(k_E, m)) = m$$

est assurée.

Cryptosystème

Définition

Un **cryptosystème** est un quintuplé $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ formé de

- un ensemble de clés $\mathcal{K}$
- un ensemble de messages clairs possibles , $\mathcal{M}$,
- un ensemble de messages chiffrés possibles $\mathcal{C}$
- un algorithme de chiffrage, représenté par une fonction $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$,
- et un procédé de déchiffrage , représenté par une fonction $D : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$.

On suppose que pour tout $m \in \mathcal{M}$ il existe une paide de clés de chiffrement et de déchiffrement telles que la relation

$$D(k_D, E(k_E, m)) = m$$

est assurée.

Cryptosystème

Définition

Un **cryptosystème** est un quintuplé $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ formé de

- un ensemble de clés $\mathcal{K}$
- un ensemble de messages clairs possibles , $\mathcal{M}$,
- un ensemble de messages chiffrés possibles $\mathcal{C}$
- un algorithme de chiffrage, représenté par une fonction $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$,
- et un procédé de déchiffrage , représenté par une fonction $D : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$.

On suppose que pour tout $m \in \mathcal{M}$ il existe une paide de clés de chiffrement et de déchiffrement telles que la relation

$$D(k_D, E(k_E, m)) = m$$

est assurée.

Cryptosystème

Définition

Un **cryptosystème** est un quintuplé $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ formé de

- un ensemble de clés $\mathcal{K}$
- un ensemble de messages clairs possibles , $\mathcal{M}$,
- un ensemble de messages chiffrés possibles $\mathcal{C}$
- un algorithme de chiffrage, représenté par une fonction $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$,
- et un procédé de déchiffrage , représenté par une fonction $D : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$.

On suppose que pour tout $m \in \mathcal{M}$ il existe une paide de clés de chiffrement et de déchiffrement telles que la relation

$$D(k_D, E(k_E, m)) = m$$

est assurée.

Cryptosystème

Définition

Un **cryptosystème** est un quintuplé $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ formé de

- un ensemble de clés $\mathcal{K}$
- un ensemble de messages clairs possibles , $\mathcal{M}$,
- un ensemble de messages chiffrés possibles $\mathcal{C}$
- un algorithme de chiffrage, représenté par une fonction $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$,
- et un procédé de déchiffrage , représenté par une fonction $D : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$.

On suppose que pour tout $m \in \mathcal{M}$ il existe une paide de clés de chiffrement et de déchiffrement telles que la relation

$$D(k_D, E(k_E, m)) = m$$

est assurée.

Cryptosystème

Définition

Un **cryptosystème** est un quintuplé $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ formé de

- un ensemble de clés $\mathcal{K}$
- un ensemble de messages clairs possibles , $\mathcal{M}$,
- un ensemble de messages chiffrés possibles $\mathcal{C}$
- un algorithme de chiffrage, représenté par une fonction $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$,
- et un procédé de déchiffrage , représenté par une fonction $D : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$.

On suppose que pour tout $m \in \mathcal{M}$ il existe une paide de clés de chiffrement et de déchiffrement telles que la relation

$$D(k_D, E(k_E, m)) = m$$

est assurée.

Cryptographie. Un peu d'histoire

- Cryptographie signifie "écriture cachée" (kryptos = caché, graphein= écriture)
- Kàma-Sutra recommandait aux femmes d'apprendre 60 arts, dont les échecs, la reliure, la tapisserie et... l'écriture secrète pour cacher leur liaisons
- L'une des premières méthodes de cryptage par substitution connue est celle de César (50 av. J.C). Chaque lettre de message à transmettre était remplacée par la lettre située dans l'alphabet trois positions plus loin.
- Par exemple, *A* est remplacé par "D", "Z" par "C".

Cryptographie. Un peu d'histoire

- Cryptographie signifie "écriture cachée" (kryptos = caché, graphein= écriture)
- Kàma-Sutra recommandait aux femmes d'apprendre 60 arts, dont les échecs, la reliure, la tapisserie et... l'écriture secrète pour cacher leur liaisons
- L'une des premières méthodes de cryptage par substitution connue est celle de César (50 av. J.C). Chaque lettre de message à transmettre était remplacée par la lettre située dans l'alphabet trois positions plus loin.
- Par exemple, *A* est remplacé par "D", "Z" par "C".

Cryptographie. Un peu d'histoire

- Cryptographie signifie "écriture cachée" (kryptos = caché, graphein= écriture)
- Kàma-Sutra recommandait aux femmes d'apprendre 60 arts, dont les échecs, la reliure, la tapisserie et... l'écriture secrète pour cacher leur liaisons
- L'une des premières méthodes de cryptage par substitution connue est celle de César (50 av. J.C). Chaque lettre de message à transmettre était remplacée par la lettre située dans l'alphabet trois positions plus loin.
- Par exemple, A est remplacé par "D", "Z" par "C".

Cryptographie. Un peu d'histoire

- Cryptographie signifie "écriture cachée" (kryptos = caché, graphein= écriture)
- Kàma-Sutra recommandait aux femmes d'apprendre 60 arts, dont les échecs, la reliure, la tapisserie et... l'écriture secrète pour cacher leur liaisons
- L'une des premières méthodes de cryptage par substitution connue est celle de César (50 av. J.C). Chaque lettre de message à transmettre était remplacée par la lettre située dans l'alphabet trois positions plus loin.
- Par exemple, A est remplacé par "D", "Z" par "C".

Cryptographie. Un peu d'histoire.

- Les méthodes de substitution, analogues à celle de César, consistent à apparier les lettres de l'alphabet et à remplacer chaque lettre de message par la lettre de sa paire
- Elles ont été massivement utilisées jusqu'à la fin du 1er millénaire
- IXème siècle : le savant arabe AL-Kindi rédige le premier traité connu sur l'analyse fréquentielle comme technique de cryptanalyse
- Il montre comment retrouver le message en analysant les fréquences d'occurrence de caractères dans une langue donnée
- Le cryptage par substitution devient trop vulnérable

Cryptographie. Un peu d'histoire.

- Les méthodes de substitution, analogues à celle de César, consistent à apparier les lettres de l'alphabet et à remplacer chaque lettre de message par la lettre de sa paire
- Elles ont été massivement utilisées jusqu'à la fin du 1er millénaire
- IXème siècle : le savant arabe AL-Kindi rédige le premier traité connu sur l'analyse fréquentielle comme technique de cryptanalyse
- Il montre comment retrouver le message en analysant les fréquences d'occurrence de caractères dans une langue donnée
- Le cryptage par substitution devient trop vulnérable

Cryptographie. Un peu d'histoire.

- Les méthodes de substitution, analogues à celle de César, consistent à apparier les lettres de l'alphabet et à remplacer chaque lettre de message par la lettre de sa paire
- Elles ont été massivement utilisées jusqu'à la fin du 1er millénaire
- IXème siècle : le savant arabe AL-Kindi rédige le premier traité connu sur l'analyse fréquentielle comme technique de cryptanalyse
- Il montre comment retrouver le message en analysant les fréquences d'occurrence de caractères dans une langue donnée
- Le cryptage par substitution devient trop vulnérable

Cryptographie. Un peu d'histoire.

- Les méthodes de substitution, analogues à celle de César, consistent à apparier les lettres de l'alphabet et à remplacer chaque lettre de message par la lettre de sa paire
- Elles ont été massivement utilisées jusqu'à la fin du 1er millénaire
- IXème siècle : le savant arabe AL-Kindi rédige le premier traité connu sur l'analyse fréquentielle comme technique de cryptanalyse
- Il montre comment retrouver le message en analysant les fréquences d'occurrence de caractères dans une langue donnée
- Le cryptage par substitution devient trop vulnérable

Cryptographie. Un peu d'histoire.

- Les méthodes de substitution, analogues à celle de César, consistent à apparier les lettres de l'alphabet et à remplacer chaque lettre de message par la lettre de sa paire
- Elles ont été massivement utilisées jusqu'à la fin du 1er millénaire
- IXème siècle : le savant arabe AL-Kindi rédige le premier traité connu sur l'analyse fréquentielle comme technique de cryptanalyse
- Il montre comment retrouver le message en analysant les fréquences d'occurrence de caractères dans une langue donnée
- Le cryptage par substitution devient trop vulnérable

Cryptographie. Notion de sécurité.

Principe de A. Kerckhoffs (fin XIXe)

La sécurité d'un cryptosystème ne doit pas reposer sur la non divulgation de la fonction de cryptage mais uniquement sur la non divulgation de la clé.

Sécurité : approche de Shannon

- 1949. Publication par C. Shannon de l'article "Communication Theory of Secrecy Systems" dans la revue Bell System Technical Journal.
- Le concept d'entropie est utilisé pour analyser et quantifier la sécurité d'un cryptosystème.

Sécurité : approche de Shannon

- 1949. Publication par C. Shannon de l'article "Communication Theory of Secrecy Systems" dans la revue Bell System Technical Journal.
- Le concept d'entropie est utilisé pour analyser et quantifier la sécurité d'un cryptosystème.

Cryptosystème parfaitement sûr

- On associe aux cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ trois variables aléatoires :
- $M \in \mathcal{M}$ représente le choix d'un message clair
- $K \in \mathcal{K}$ représente le choix d'une clé
- $C \in \mathcal{C}$ représente le choix d'un chiffré
- on suppose que le message clair et la clé sont choisis de façon indépendante

Cryptosystème parfaitement sûr

- On associe aux cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ trois variables aléatoires :
- $M \in \mathcal{M}$ représente le choix d'un message clair
- $K \in \mathcal{K}$ représente le choix d'une clé
- $C \in \mathcal{C}$ représente le choix d'un chiffré
- on suppose que le message clair et la clé sont choisis de façon indépendante

Cryptosystème parfaitement sûr

- On associe aux cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ trois variables aléatoires :
- $M \in \mathcal{M}$ représente le choix d'un message clair
- $K \in \mathcal{K}$ représente le choix d'une clé
- $C \in \mathcal{C}$ représente le choix d'un chiffré
- on suppose que le message clair et la clé sont choisis de façon indépendante

Cryptosystème parfaitement sûr

- On associe aux cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ trois variables aléatoires :
- $M \in \mathcal{M}$ représente le choix d'un message clair
- $K \in \mathcal{K}$ représente le choix d'une clé
- $C \in \mathcal{C}$ représente le choix d'un chiffré
- on suppose que le message clair et la clé sont choisis de façon indépendante

Cryptosystème parfaitement sûr

- On associe aux cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$ trois variables aléatoires :
- $M \in \mathcal{M}$ représente le choix d'un message clair
- $K \in \mathcal{K}$ représente le choix d'une clé
- $C \in \mathcal{C}$ représente le choix d'un chiffré
- on suppose que le message clair et la clé sont choisis de façon indépendante

Cryptosystème parfaitement sûr

Définition

Soit un cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$. Soient M et C les variables aléatoires représentant le choix d'un message claire et d'un chiffré. Le système est dit parfaitement sûr ssi

$$H(M|C) = H(M)$$

La connaissance du chiffré n'apporte aucune information sur le message clair.

Cryptosystème parfaitement sûr

Définition

Soit un cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$. Soient M et C les variables aléatoires représentant le choix d'un message claire et d'un chiffré. Le système est dit parfaitement sûr ssi

$$H(M|C) = H(M)$$

La connaissance du chiffré n'apporte aucune information sur le message clair.

Exemple de chiffrement parfaitement sûr : le chiffre de Vernam

- Cette méthode a été proposée en 1917.
- Les messages clairs sont des suites de bits de longueur n. L'espace des messages est alors $\mathcal{M} = 0,1^n$.
- Les clés sont les suites binaires de même longueur que les messages : $\mathcal{K} = \mathcal{M} = 0,1^n$.
- La fonction de chiffrement : pour tout $m = m_1 \dots m_n$ et pour tout $k = k_1 \dots k_n$

$$c = E(k, m) = k \oplus m = m_1 \oplus k_1 \dots m_n \oplus k_n$$

Exemple de chiffrement parfaitement sûr : le chiffre de Vernam

- Cette méthode a été proposée en 1917.
- Les messages clairs sont des suites de bits de longueur n. L'espace des messages est alors $\mathcal{M} = 0, 1^n$.
- Les clés sont les suites binaires de même longueur que les messages : $\mathcal{K} = \mathcal{M} = 0, 1^n$.
- La fonction de chiffrement : pour tout $m = m_1 \dots m_n$ et pour tout $k = k_1 \dots k_n$

$$c = E(k, m) = k \oplus m = m_1 \oplus k_1 \dots m_n \oplus k_n$$

Exemple de chiffrement parfaitement sûr : le chiffre de Vernam

- Cette méthode a été proposée en 1917.
- Les messages clairs sont des suites de bits de longueur n. L'espace des messages est alors $\mathcal{M} = 0,1^n$.
- Les clés sont les suites binaires de même longueur que les messages : $\mathcal{K} = \mathcal{M} = 0,1^n$.
- La fonction de chiffrement : pour tout $m = m_1 \dots m_n$ et pour tout $k = k_1 \dots k_n$

$$c = E(k, m) = k \oplus m = m_1 \oplus k_1 \dots m_n \oplus k_n$$

Exemple de chiffrement parfaitement sûr : le chiffre de Vernam

- Cette méthode a été proposée en 1917.
- Les messages clairs sont des suites de bits de longueur n. L'espace des messages est alors $\mathcal{M} = 0,1^n$.
- Les clés sont les suites binaires de même longueur que les messages : $\mathcal{K} = \mathcal{M} = 0,1^n$.
- La fonction de chiffrement : pour tout $m = m_1 \dots m_n$ et pour tout $k = k_1 \dots k_n$

$$c = E(k, m) = k \oplus m = m_1 \oplus k_1 \dots m_n \oplus k_n$$

Exemple de chiffrement parfaitement sûr : le chiffre de Vernam

Proposition

Le chiffrement de Vernam est parfaitement sur

Théorème

Dans un système cryptographique parfaitement sûr on a

$$H(K) \geq H(M)$$

En particulier, si tous les messages et toutes les clés sont équiprobables, les clés sont de longueur au moins égale à celle des messages.

Exemple de chiffrement parfaitement sûr : le chiffre de Vernam

Proposition

Le chiffrement de Vernam est parfaitement sur

Théorème

Dans un système cryptographique parfaitement sûr on a

$$H(K) \geq H(M)$$

En particulier, si tous les messages et toutes les clés sont équiprobables, les clés sont de longueur au moins égale à celle des messages.

Attention ! Un chiffrement parfaitement sûr n'est pas invulnérable !

- Dans le cas de chiffrement de Vernam il suffit d'un bloc de message clair pour découvrir la clé :

$$k = c \oplus m$$

- Il est alors nécessaire de changer de clé à chaque bloc ; d'où le nom de "masque jetable" ;
- En pratique, cette procédure et très lourde : les clés ont la même longueur que les messages
- Ce chiffrement a été utilisé pour le téléphone rouge : la clé a été envoyé par porteur

Attention ! Un chiffrement parfaitement sûr n'est pas invulnérable !

- Dans le cas de chiffrement de Vernam il suffit d'un bloc de message clair pour découvrir la clé :

$$k = c \oplus m$$

- Il est alors nécessaire de changer de clé à chaque bloc ; d'où le nom de "masque jetable" ;
- En pratique, cette procédure et très lourde : les clés ont la même longueur que les messages
- Ce chiffrement a été utilisé pour le téléphone rouge : la clé a été envoyé par porteur

Attention ! Un chiffrement parfaitement sûr n'est pas invulnérable !

- Dans le cas de chiffrement de Vernam il suffit d'un bloc de message clair pour découvrir la clé :

$$k = c \oplus m$$

- Il est alors nécessaire de changer de clé à chaque bloc ; d'où le nom de "masque jetable" ;
- En pratique, cette procédure et très lourde : les clés ont la même longueur que les messages
- Ce chiffrement a été utilisé pour le téléphone rouge : la clé a été envoyé par porteur

Attention ! Un chiffrement parfaitement sûr n'est pas invulnérable !

- Dans le cas de chiffrement de Vernam il suffit d'un bloc de message clair pour découvrir la clé :

$$k = c \oplus m$$

- Il est alors nécessaire de changer de clé à chaque bloc ; d'où le nom de "masque jetable" ;
- En pratique, cette procédure et très lourde : les clés ont la même longueur que les messages
- Ce chiffrement a été utilisé pour le téléphone rouge : la clé a été envoyé par porteur

Et si les clés sont plus courtes ? Quelques définitions.

- Soit un langage composé de mots de longueur donnée N sur un alphabet donné A. On associe la variable aléatoire M au choix au hasard d'un mot du langage.
- Taux du langage (ou taux d'entropie)

$$r = \lim_{N \to \infty} \frac{H(M)}{N}$$

représente la quantité moyenne d'information par caractère de message
- Si le nombre de caractères de l'alphabet est L on appelle taux maximal du langage la quantité

$$R = log_2 L$$

Il s'agit de l'entropie maximale d'un caractère
- Enfin on appelle redondance du langage la différence

$$D = R - r$$

Et si les clés sont plus courtes ? Quelques définitions.

- Soit un langage composé de mots de longueur donnée N sur un alphabet donné A. On associe la variable aléatoire M au choix au hasard d'un mot du langage.
- Taux du langage (ou taux d'entropie)

$$r = \lim_{N \to \infty} \frac{H(M)}{N}$$

représente la quantité moyenne d'information par caractère de message

- Si le nombre de caractères de l'alphabet est L on appelle taux maximal du langage la quantité

$$R = log_2 L$$

Il s'agit de l'entropie maximale d'un caractère

- Enfin on appelle redondance du langage la différence

$$D = R - r$$

Et si les clés sont plus courtes ? Quelques définitions.

- Soit un langage composé de mots de longueur donnée N sur un alphabet donné A. On associe la variable aléatoire M au choix au hasard d'un mot du langage.
- Taux du langage (ou taux d'entropie)

$$r = \lim_{N \to \infty} \frac{H(M)}{N}$$

  représente la quantité moyenne d'information par caractère de message
- Si le nombre de caractères de l'alphabet est L on appelle taux maximal du langage la quantité

$$R = log_2 L$$

  Il s'agit de l'entropie maximale d'un caractère
- Enfin on appelle redondance du langage la différence

$$D = R - r$$

Et si les clés sont plus courtes ? Quelques définitions.

- Soit un langage composé de mots de longueur donnée N sur un alphabet donné A. On associe la variable aléatoire M au choix au hasard d'un mot du langage.
- Taux du langage (ou taux d'entropie)

$$r = \lim_{N \to \infty} \frac{H(M)}{N}$$

  représente la quantité moyenne d'information par caractère de message
- Si le nombre de caractères de l'alphabet est L on appelle taux maximal du langage la quantité

$$R = log_2 L$$

  Il s'agit de l'entropie maximale d'un caractère
- Enfin on appelle redondance du langage la différence

$$D = R - r$$

Distance d'unicité

Définition

Soit un cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$. On appelle la distance d'unicité le plus petit nombre de messages chiffrés dont il est nécessaire disposer pour que l'incertitude résiduelle sur la clé sachant ces chiffrés soit nulle :

$$n_0 : \ H(K|C_1, \ldots, C_{n_0}) = 0$$

Attention

Il faut interpréter cette définitoion dans le sens suivant : si l'on ne dispose pas de longueur suffisante de message chiffré donnée par la distance d'unicité, il est impossible de déterminer la clé avec certitude.
Ce résultat ne permet surtout pas de se prononcer sur la puissance de calcul nécessaire pour découvrir la clé ni sur les moyens d'y parvenir.

Distance d'unicité

Définition

Soit un cryptosystème $(\mathcal{K}, \mathcal{M}, \mathcal{C}, E, D)$. On appelle la distance d'unicité le plus petit nombre de messages chiffrés dont il est nécessaire disposer pour que l'incertitude résiduelle sur la clé sachant ces chiffrés soit nulle :

$$n_0 : \ H(K|C_1, \ldots, C_{n_0}) = 0$$

Attention

Il faut interpréter cette définitoion dans le sens suivant : si l'on ne dispose pas de longueur suffisante de message chiffré donnée par la distance d'unicité, il est impossible de déterminer la clé avec certitude.
Ce résultat ne permet surtout pas de se prononcer sur la puissance de calcul nécessaire pour découvrir la clé ni sur les moyens d'y parvenir.

Distance d'unicité

Proposition

La distance d'unicité d d'un cryptosystème est égale à

$$d = \frac{H(K)}{D} = \frac{H(K)}{R - r}$$

où D est la redondance du langage.

Attention

La distance d'unicité est inversement proportionnelle à la redondance des messages clairs.
Pour améliorer la distance d'unicité, il est préférable de traiter en amont les messages de façon à réduire la redondance. Par exemple, une compression sans pertes par codage de Huffman ou autre peut être utilisée avant le chiffrement.

Distance d'unicité

Proposition

La distance d'unicité d d'un cryptosystème est égale à

$$d = \frac{H(K)}{D} = \frac{H(K)}{R - r}$$

où D est la redondance du langage.

Attention

La distance d'unicité est inversement proportionnelle à la redondance des messages clairs.
Pour améliorer la distance d'unicité, il est préférable de traiter en amont les messages de façon à réduire la redondance. Par exemple, une compression sans pertes par codage de Huffman ou autre peut être utilisée avant le chiffrement.

Distance d'unicité : exemple

- Considérons un chiffrement par substitution.
- La clé est alors une permutation quelconque de l'alphabet latin σ.
- Chaque lettre m_i d'un message $m = m_1, \ldots, m_n$ est remplacée par $\sigma(m_i)$
- L'ensemble des clés est alors l'ensemble de toutes les permutations possibles
- $H(K) = \log_2(26!)$
- Pour le français on a $r \simeq 3.97$, $R = 4.67$. Ainsi la redondance est $D = 0.7$.
- on obtient alors la distance d'unicité d'un chiffrement par substitution :

$$d = \frac{H(K)}{D} = \frac{log_2(26!)}{0.7} \simeq 126$$

Distance d'unicité : exemple

- Considérons un chiffrement par substitution.
- La clé est alors une permutation quelconque de l'alphabet latin σ.
- Chaque lettre m_i d'un message $m = m_1, \ldots, m_n$ est remplacée par $\sigma(m_i)$
- L'ensemble des clés est alors l'ensemble de toutes les permutations possibles
- $H(K) = \log_2(26!)$
- Pour le français on a $r \simeq 3.97$, $R = 4.67$. Ainsi la redondance est $D = 0.7$.
- on obtient alors la distance d'unicité d'un chiffrement par substitution :

$$d = \frac{H(K)}{D} = \frac{log_2(26!)}{0.7} \simeq 126$$

Distance d'unicité : exemple

- Considérons un chiffrement par substitution.
- La clé est alors une permutation quelconque de l'alphabet latin σ.
- Chaque lettre m_i d'un message $m = m_1, \ldots, m_n$ est remplacée par $\sigma(m_i)$
- L'ensemble des clés est alors l'ensemble de toutes les permutations possibles
- $H(K) = \log_2(26!)$
- Pour le français on a $r \simeq 3.97$, $R = 4.67$. Ainsi la redondance est $D = 0.7$.
- on obtient alors la distance d'unicité d'un chiffrement par substitution :

$$d = \frac{H(K)}{D} = \frac{log_2(26!)}{0.7} \simeq 126$$

Distance d'unicité : exemple

- Considérons un chiffrement par substitution.
- La clé est alors une permutation quelconque de l'alphabet latin σ.
- Chaque lettre m_i d'un message $m = m_1, \ldots, m_n$ est remplacée par $\sigma(m_i)$
- L'ensemble des clés est alors l'ensemble de toutes les permutations possibles
- $H(K) = \log_2(26!)$
- Pour le français on a $r \simeq 3.97$, $R = 4.67$. Ainsi la redondance est $D = 0.7$.
- on obtient alors la distance d'unicité d'un chiffrement par substitution :

$$d = \frac{H(K)}{D} = \frac{log_2(26!)}{0.7} \simeq 126$$

Distance d'unicité : exemple

- Considérons un chiffrement par substitution.
- La clé est alors une permutation quelconque de l'alphabet latin σ.
- Chaque lettre m_i d'un message $m = m_1, \ldots, m_n$ est remplacée par $\sigma(m_i)$
- L'ensemble des clés est alors l'ensemble de toutes les permutations possibles
- $H(K) = \log_2(26!)$
- Pour le français on a $r \simeq 3.97$, $R = 4.67$. Ainsi la redondance est $D = 0.7$.
- on obtient alors la distance d'unicité d'un chiffrement par substitution :

$$d = \frac{H(K)}{D} = \frac{log_2(26!)}{0.7} \simeq 126$$

Distance d'unicité : exemple

- Considérons un chiffrement par substitution.
- La clé est alors une permutation quelconque de l'alphabet latin σ.
- Chaque lettre m_i d'un message $m = m_1, \ldots, m_n$ est remplacée par $\sigma(m_i)$
- L'ensemble des clés est alors l'ensemble de toutes les permutations possibles
- $H(K) = \log_2(26!)$
- Pour le français on a $r \simeq 3.97$, $R = 4.67$. Ainsi la redondance est $D = 0.7$.
- on obtient alors la distance d'unicité d'un chiffrement par substitution :

$$d = \frac{H(K)}{D} = \frac{log_2(26!)}{0.7} \simeq 126$$

Distance d'unicité : exemple

- Considérons un chiffrement par substitution.
- La clé est alors une permutation quelconque de l'alphabet latin σ.
- Chaque lettre m_i d'un message $m = m_1, \ldots, m_n$ est remplacée par $\sigma(m_i)$
- L'ensemble des clés est alors l'ensemble de toutes les permutations possibles
- $H(K) = \log_2(26!)$
- Pour le français on a $r \simeq 3.97$, $R = 4.67$. Ainsi la redondance est $D = 0.7$.
- on obtient alors la distance d'unicité d'un chiffrement par substitution :

$$d = \frac{H(K)}{D} = \frac{log_2(26!)}{0.7} \simeq 126$$