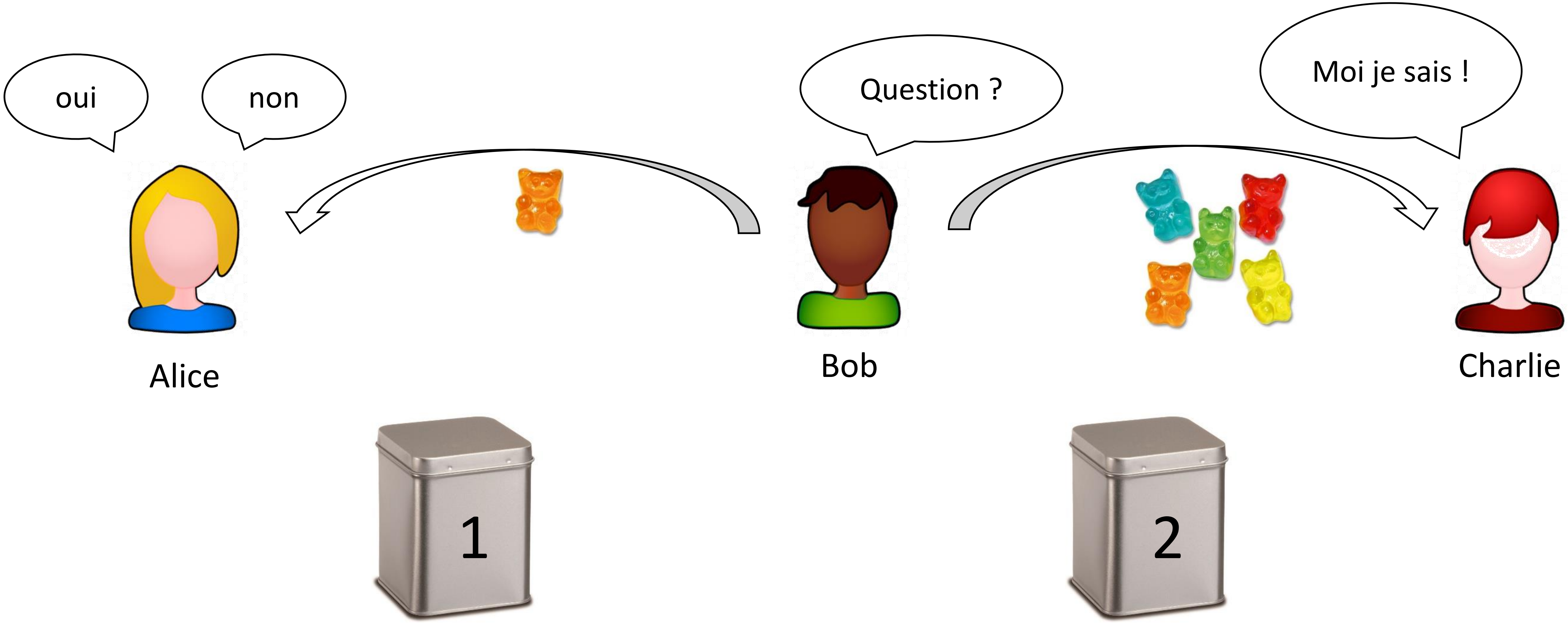
Mesurer l'information

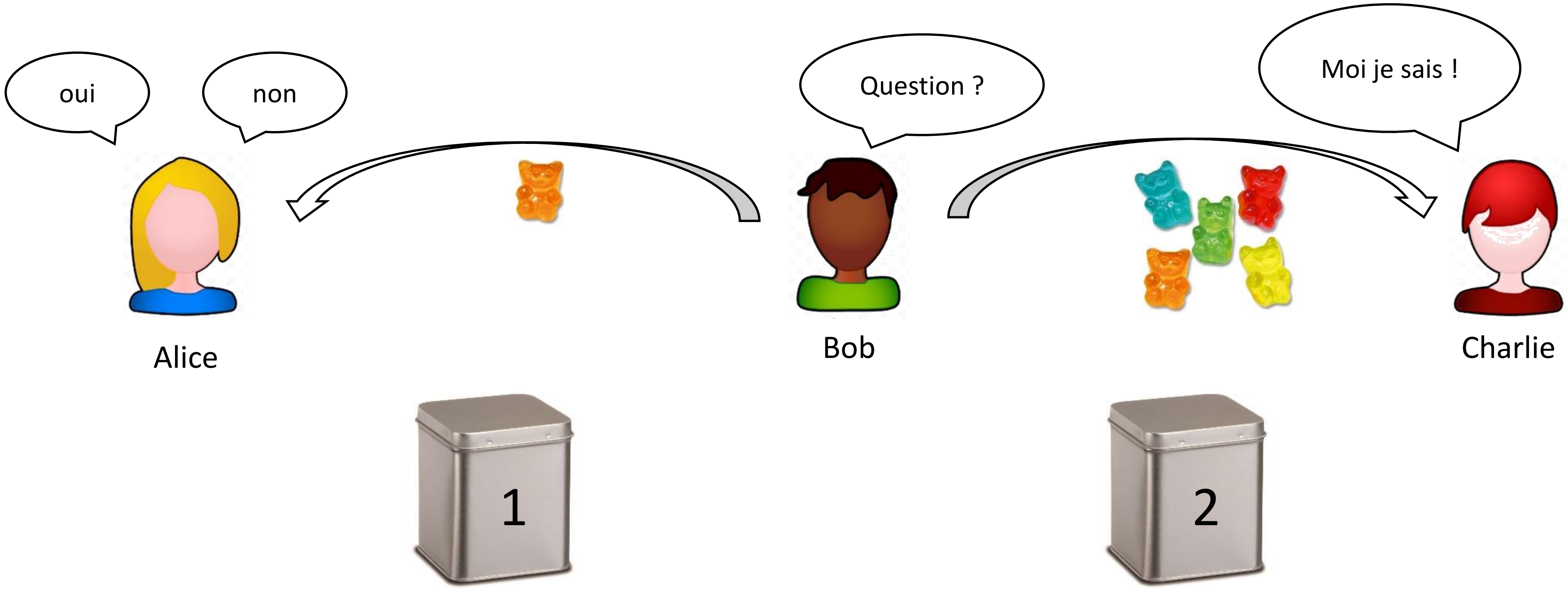
Mesure de la quantité d'information

1. Cas simple :



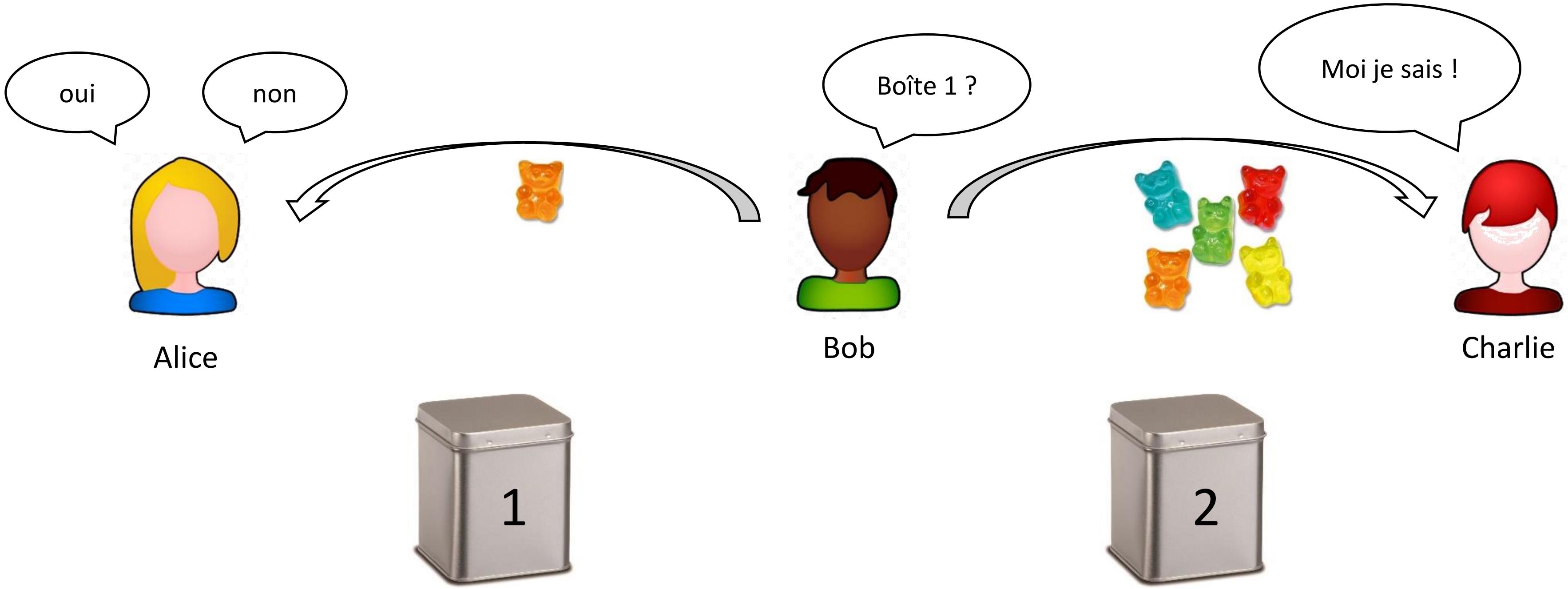
Mesure de la quantité d'information

- Combien de questions sont nécessaires pour trouver la bonne boîte ?
- Quel est le coût envers Alice ?
- Combien, au plus, doit-on offrir à Charlie ?

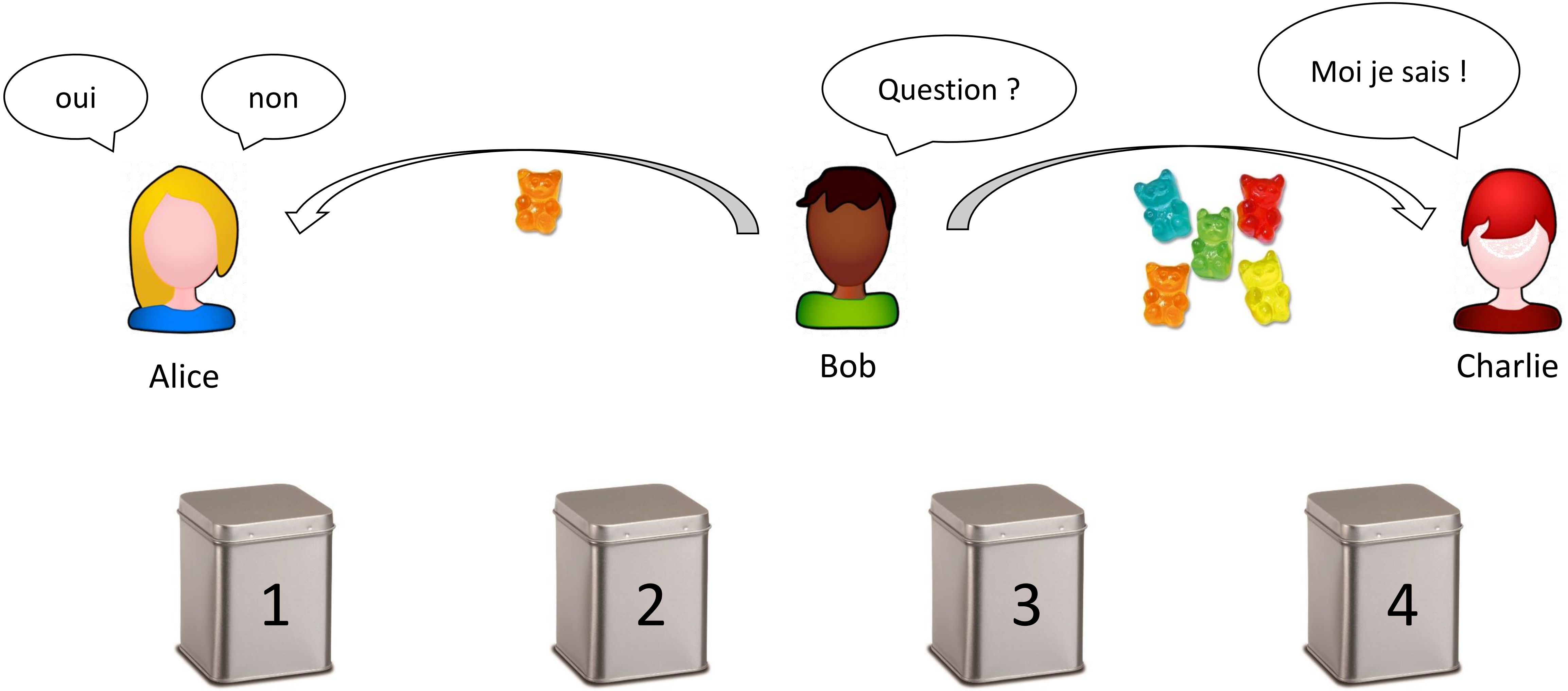


Mesure de la quantité d'information

- Combien de questions sont nécessaires pour trouver la bonne boîte ? 1
- Quel est le coût envers Alice ? 1 ourson
- Combien, au plus, doit-on offrir à Charlie ? 1 ourson

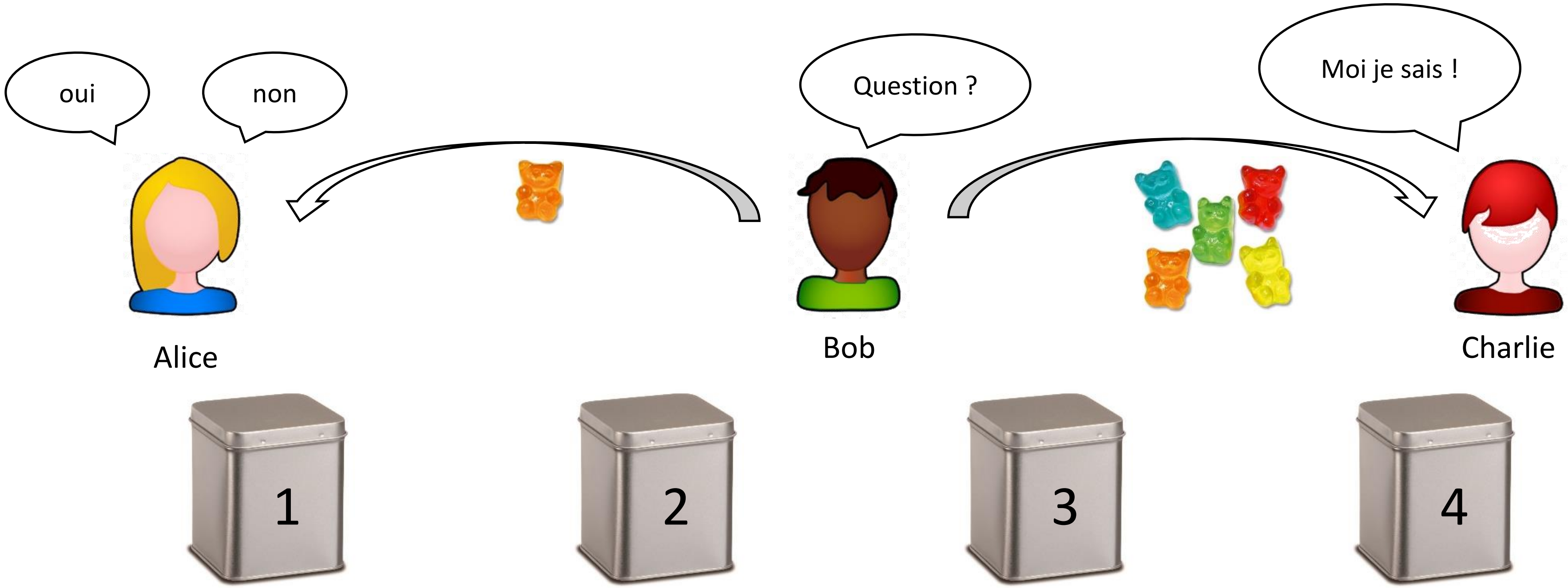


Mesure de la quantité d'information



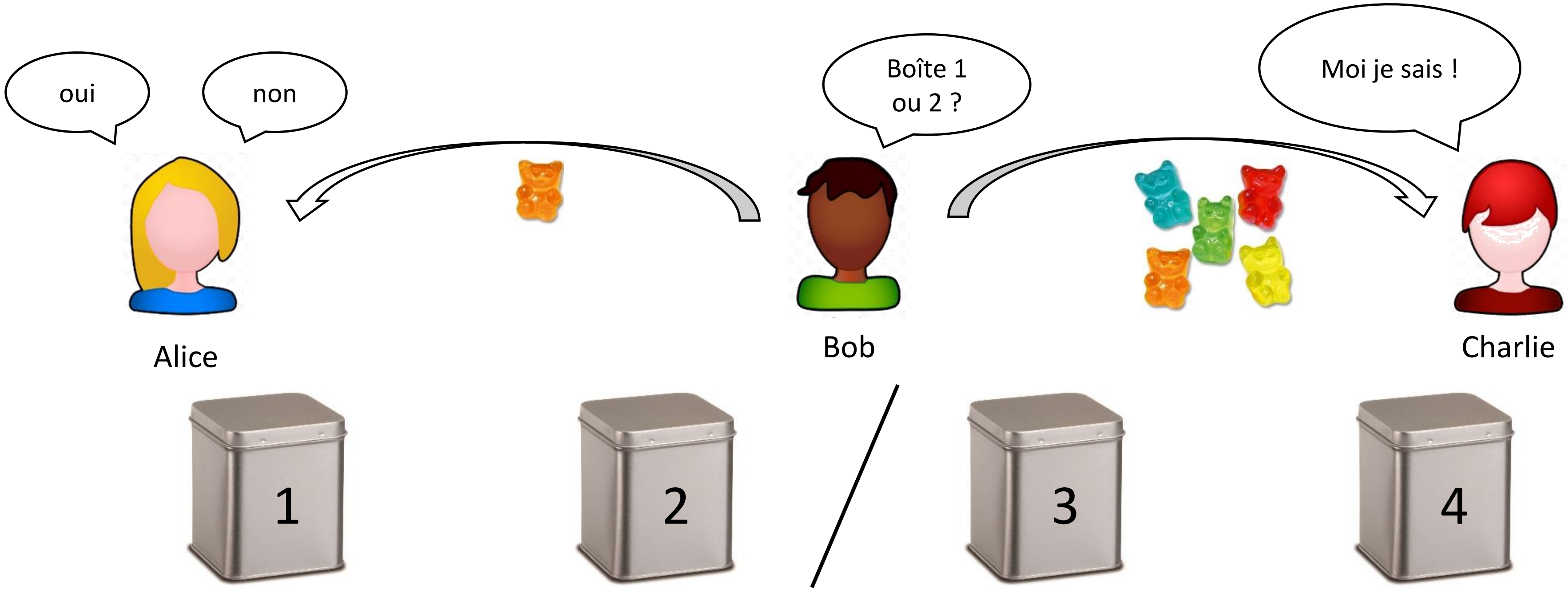
Mesure de la quantité d'information

- Combien de questions, en moyenne, pour trouver la bonne boîte ?
- Quel est le coût moyen envers Alice ?
- Combien, au plus, doit-on offrir à Charlie ?

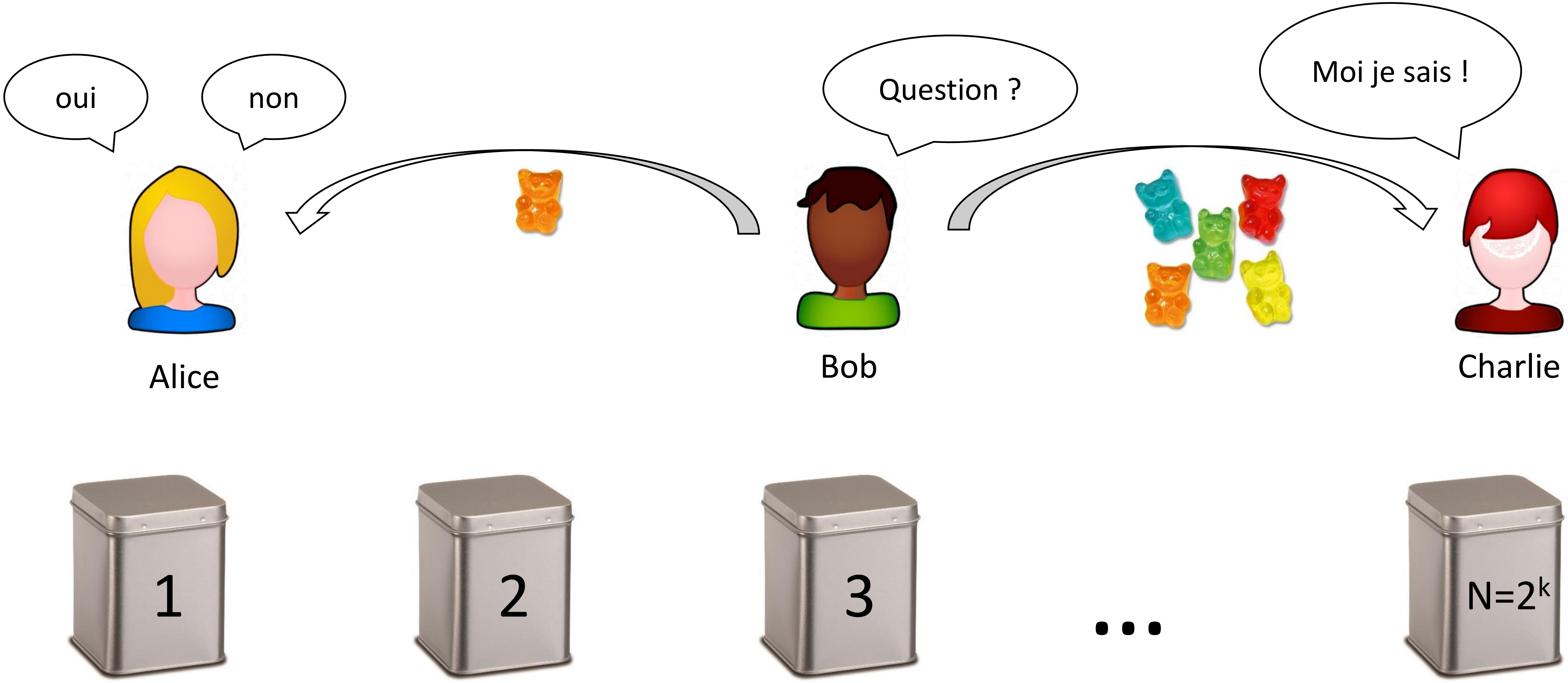


Mesure de la quantité d'information

- Combien de questions, en moyenne, pour trouver la bonne boîte ? 2
- Quel est le coût moyen envers Alice ? 2 oursons
- Combien, au plus, doit-on offrir à Charlie ? 2 oursons



Mesure de la quantité d'information



Mesure de la quantité d'information

Astuce : numérotation en base binaire

111000011101001101000101 0……..100011101010000110 : k chiffres au plus

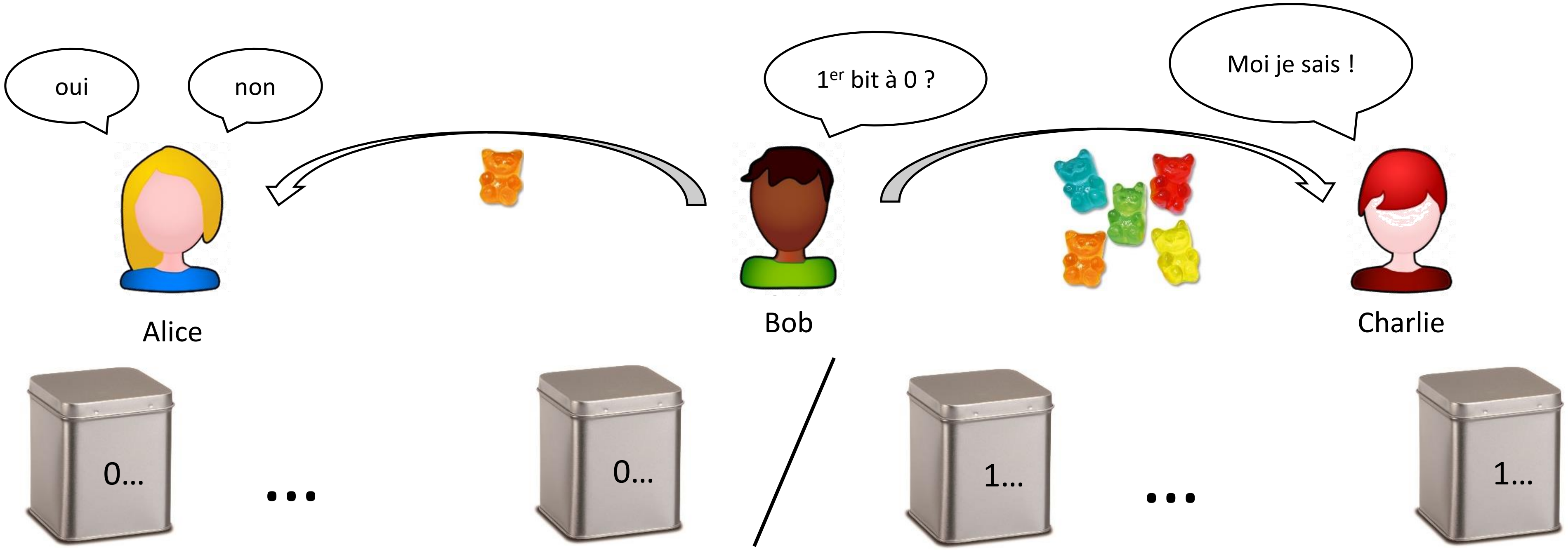
Recherche dichotomique : *« Le $i^{ème}$ chiffre est-il un 0 ? »*

Un **bit** est égal à la quantité d'information fournie par le choix d'une alternative parmi deux équiprobables.

bit = binary digit = symbole binaire

Mesure de la quantité d'information

- Combien de questions, en moyenne, pour trouver la bonne boîte ? k
- Quel est le coût moyen envers Alice ? k oursons
- Combien, au plus, doit-on offrir à Charlie ? k oursons = $\log_2(N)$



Mesure de la quantité d'information

Conclusions :

- L’information détenue par Charlie a un **certain prix**
- Ce prix est lié à la **quantité** d’information que Charlie détient :
  - Si la boîte est facile à trouver, alors Charlie détient une connaissance *pas* très impressionnante : peu d’information
  - Si la boîte est difficile à trouver, alors Charlie détient beaucoup de connaissance / d’information

Quantité d'information *relative à un évènement*

Il y a *n* boîtes bleues.

Quantité d’information relative à un évènement

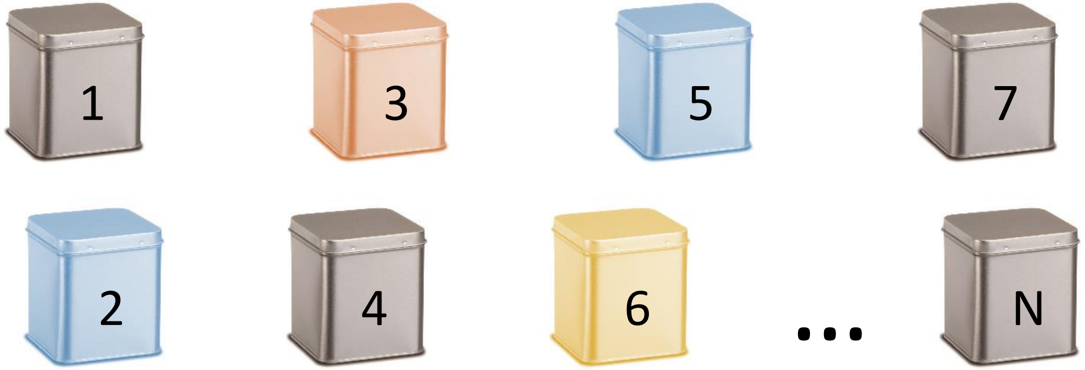
Valeur de l’information « la boîte est bleue » ?

- Prix pour trouver la boîte sans l’information : $\log_2(N)$
- Prix pour trouver la boîte avec l’information : $\log_2(n)$

(il y a *n* boîtes bleues)

→ $\log_2(N) - \log_2(n) = \log_2(N/n)$

Charlie



Quantité d'information relative à un évènement

Définition formelle :

On définit la **quantité d'information** comme une fonction croissante de N/n avec :

- N le nombre d'évènements possibles
- *n* le nombre d'évènements du sous-ensemble délimité par l'information

Mesure proposée par Shannon : $I = \log_2(N/n)$, appelée « bit » pour « binary digit »

Rappel:
k oursons = $\log_2(N)$

Quantité d'information relative à un évènement

Mesure de Shannon : $\mathbb{I} = \log_2(\frac{N}{n})$

Remarque :

Un évènement peu probable contient plus d’information qu’un évènement banal

- « Il fait 8°C en août à Toulon. » vs « Il fait 27°C en août à Toulon. »
- Pour *n* petit, $\mathbb{I}$ est grand

Quantité d'information *moyenne*

Imaginons que l'espace des boîtes est divisé en sous-ensembles distincts :

- n_{bleue} boîtes bleues
- n_{grise} boîtes grises
- n_{jaune} boîtes jaunes
- n_{orange} boîtes oranges
- n_{verte} boîtes vertes

avec $n_{\text{bleue}} + n_{\text{grise}} + n_{\text{jaune}} + n_{\text{orange}} + n_{\text{verte}} = N$

Quantité d'information moyenne

- L'**information** « la boîte est bleue » **vaut** $\log_2(\frac{N}{n_{bleue}})$
- La **probabilité** pour qu'une boîte soit bleue est $\frac{n_{bleue}}{N}$

- L'**information** « la boîte est grise » **vaut** $\log_2(\frac{N}{n_{grise}})$
- La **probabilité** pour qu'une boîte soit grise est $\frac{n_{grise}}{N}$

Quantité d'information moyenne

- L'**information** « la boîte est de couleur c_i » **vaut** $\log_2(\frac{N}{n_i})$
- La **probabilité** pour qu'une boîte soit de couleur c_i est $\frac{n_i}{N}$

La **valeur moyenne** de l'**information** sur la couleur est donc :

$$\frac{n_{bleue}}{N}\log_2(\frac{N}{n_{bleue}}) + \ldots + \frac{n_{grise}}{N}\log_2(\frac{N}{n_{grise}})$$

avec $\frac{n_{bleue}}{N} + \ldots + \frac{n_{grise}}{N} = 1$

Quantité d'information moyenne

- L'**information** « la boîte est de couleur c_i » **vaut** $\log_2(\frac{N}{n_i})$
- La **probabilité** pour qu'une boîte soit de couleur c_i est $\frac{n_i}{N}$

La **valeur moyenne** de l'**information** sur la couleur est donc :

$$\frac{n_{bleue}}{N}\log_2\left(\frac{N}{n_{bleue}}\right) + \ldots + \frac{n_{grise}}{N}\log_2\left(\frac{N}{n_{grise}}\right)$$

Entropie !

avec $\frac{n_{bleue}}{N} + \ldots + \frac{n_{grise}}{N} = 1$

Entropie de la distribution de probabilité

$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

avec $p_i = \frac{n_i}{N}$

L'**entropie** mesure la quantité moyenne d'information d'un ensemble d'évènements, et son incertitude.

$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

Calcul de l’entropie

1. Identifier la variable aléatoire, les valeurs qu’elle peut prendre, et sa distribution de probabilité

Exemple :

Pour calculer l’entropie d’un texte

lettre	%	*Lettre*	%	*Lettre*	%	*Lettre*	%	*Lettre*	%	*Lettre*	%
A	6,4	B	0,64	C	2,59	D	2,6	E	14,86	G	0,83
H	0,61	I	5,91	J	0,23	K	0,01	L	4,65	M	2,45
N	6,23	O	4,59	P	2,56	Q	0,81	R	5,55	S	6,97
T	5,72	U	5,06	W	0	X	0,31	Y	0,21	Z	0,08
Espace	18,35										
F	1,12	V	0,66								

Calcul de l’entropie

1. Identifier la variable aléatoire, les valeurs qu’elle peut prendre, et sa distribution de probabilité

2. Appliquer la formule avec les p_i ainsi trouvés

$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

Calcul de l'entropie

- Cas simple : variable aléatoire binaire $\mathrm{X} \in \{a, b\}$

$p_X(a) = q$ $\qquad$ $p_X(b) = 1 - q$

$H(X) = ?$

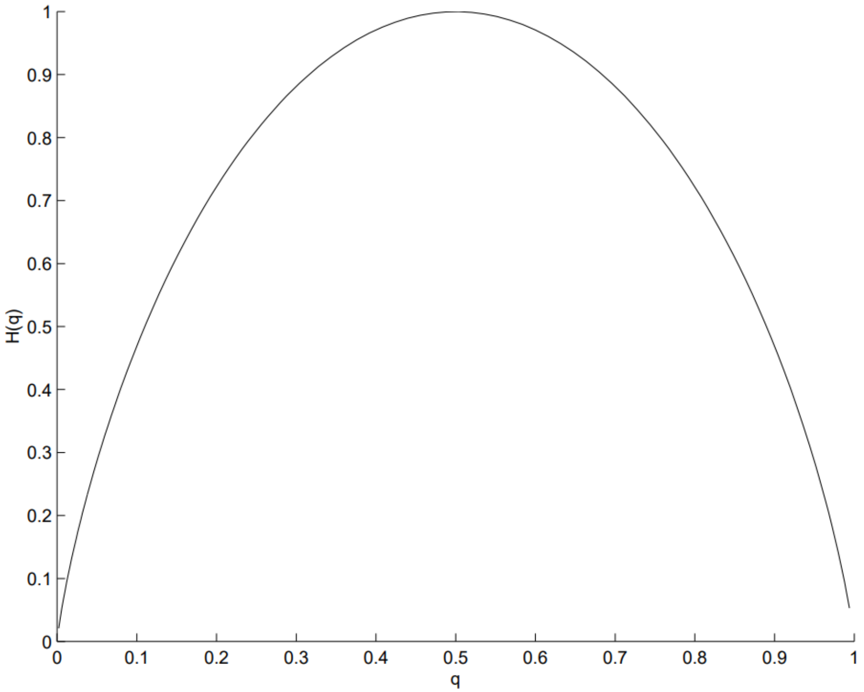
$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

Calcul de l'entropie

- Cas simple : variable aléatoire binaire $\mathrm{X} \in \{a, b\}$

$p_X(a) = q$ $\qquad$ $p_X(b) = 1 - q$

$$H(X) = -p_a \log_2 p_a - p_b \log_2 p_b$$
$$H(X) = -q \log_2 q - (1-q) \log_2(1-q)$$



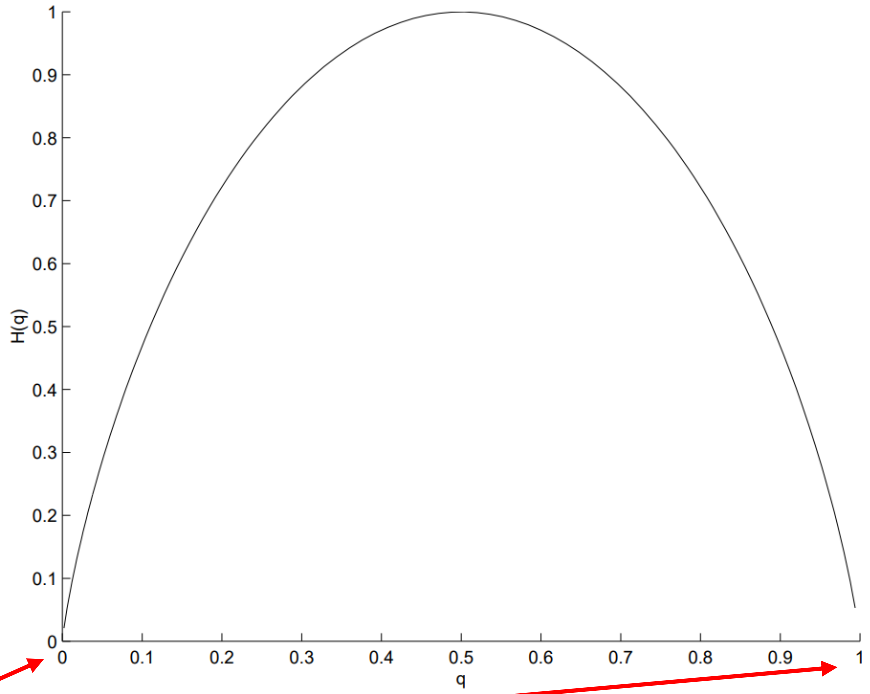
Calcul de l'entropie

- Cas simple : variable aléatoire binaire $\mathrm{X} \in \{a, b\}$

$p_X(a) = q$ $\qquad\qquad$ $p_X(b) = 1 - q$

$$H(X) = -p_a \log_2 p_a - p_b \log_2 p_b$$
$$H(X) = -q \log_2 q - (1-q) \log_2 (1-q)$$

a et b équiprobables, incertitude maximale



Cas déterministes

$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

Calcul de l'entropie

- Deuxième cas simple : variable aléatoire ternaire $\mathrm{X} \in \{a, b, c\}$ avec

$p_a + p_b + p_c = 1$

$H(X) = ?$

Quand l'entropie est-elle maximale ? Quelle est sa valeur max ?

Calcul de l'entropie

- Deuxième cas simple : variable aléatoire ternaire X $\in \{a, b, c\}$ avec

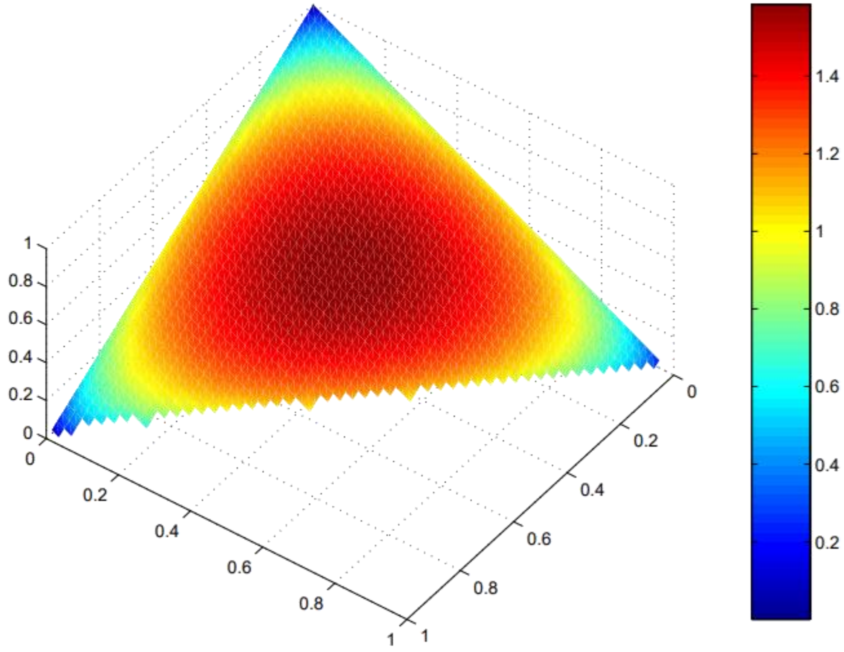
$$p_a + p_b + p_c = 1$$

$$H(X) = -p_a \log_2 p_a - p_b \log_2 p_b - p_c \log_2 p_c$$

Entropie maximale pour les valeurs

équiprobables : $p_a = p_b = p_c = \frac{1}{3}$

avec $H(X) = \log_2 3$



$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

Calcul de l'entropie

- Troisième cas simple : variable aléatoire (discrète) uniforme $\mathrm{X} \in \mathcal{X}$ avec $|\mathcal{X}| = m$ et $p(x) = \frac{1}{m}, \forall x \in \mathcal{X}$:

$H(X) = ?$

Calcul de l’entropie

- Troisième cas simple : variable aléatoire (discrète) uniforme $\mathrm{X} \in \mathcal{X}$ avec $|\mathcal{X}| = m$ et $p(x) = \frac{1}{m}, \forall x \in \mathcal{X}$:

$$H(X) = \sum_x \frac{1}{m} \log_2 m = \log_2 m$$

Calcul de l'entropie

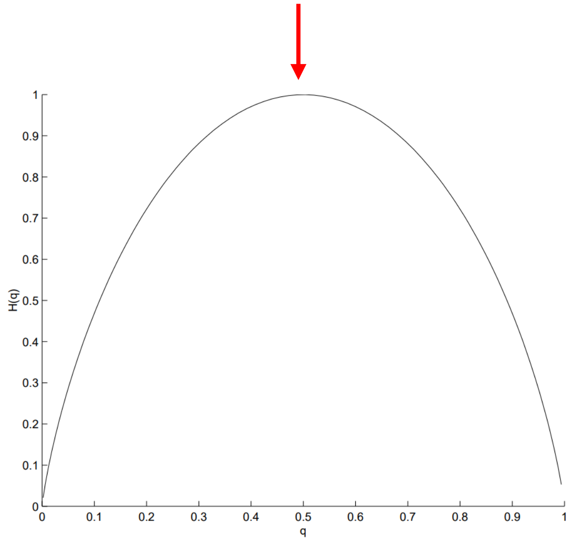
- Troisième cas simple : variable aléatoire (discrète) uniforme $\mathrm{X} \in \mathcal{X}$ avec $|\mathcal{X}| = m$ et $p(x) = \frac{1}{m}, \forall x \in \mathcal{X}$:

$$H(X) = \sum_x \frac{1}{m} \log_2 m = \log_2 m$$

Remarque : la variable binaire est un cas particulier

avec $m = 2$ pour $q = \frac{1}{2}$:

$H(X) = \log_2 2 = 1$



Calcul de l'entropie

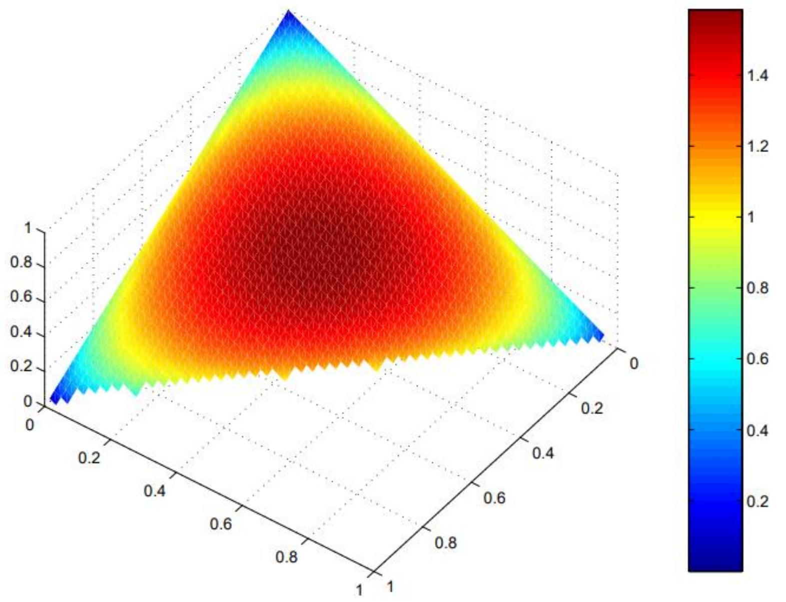
- Troisième cas simple : variable aléatoire (discrète) uniforme $\mathrm{X} \in \mathcal{X}$ avec $|\mathcal{X}| = m$ et $p(x) = \frac{1}{m}, \forall x \in \mathcal{X}$:

$$H(X) = \sum_x \frac{1}{m} \log_2 m = \log_2 m$$

Remarque : la variable ternaire est aussi

un cas particulier avec $m = 3$ pour $p_i = \frac{1}{3}$:

$$H(X) = \log_2 3 \approx 1.5850$$



Borne supérieure de l'entropie

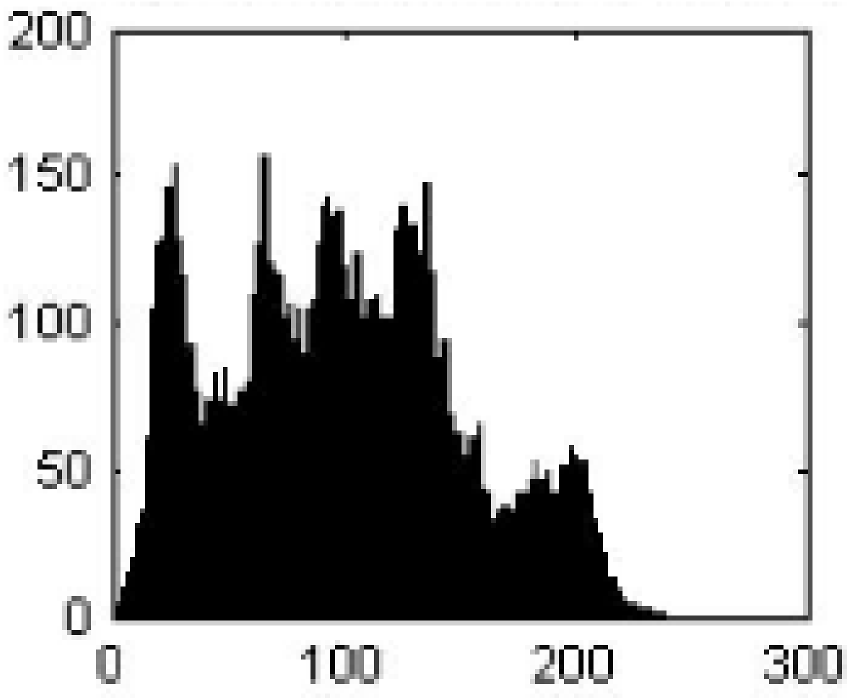
Pour des variables aléatoires discrètes $\mathrm{X} \in \mathcal{X}$ avec $|\mathcal{X}| = m$:

$\log_2 m$ est une borne supérieure de l'entropie

$$H(X) \leq \log_2 m$$

Calcul de l'entropie d'une image

Les pixels d'une image en niveaux de gris ont une valeur $i \in [0{,}255]$.



$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

Calcul de l'entropie d'une image

- Quelle est l'entropie d'une image dont tous les pixels ont la valeur 0 ?

Calcul de l’entropie d’une image

- Quelle est l’entropie d’une image dont tous les pixels ont la valeur 0 ?

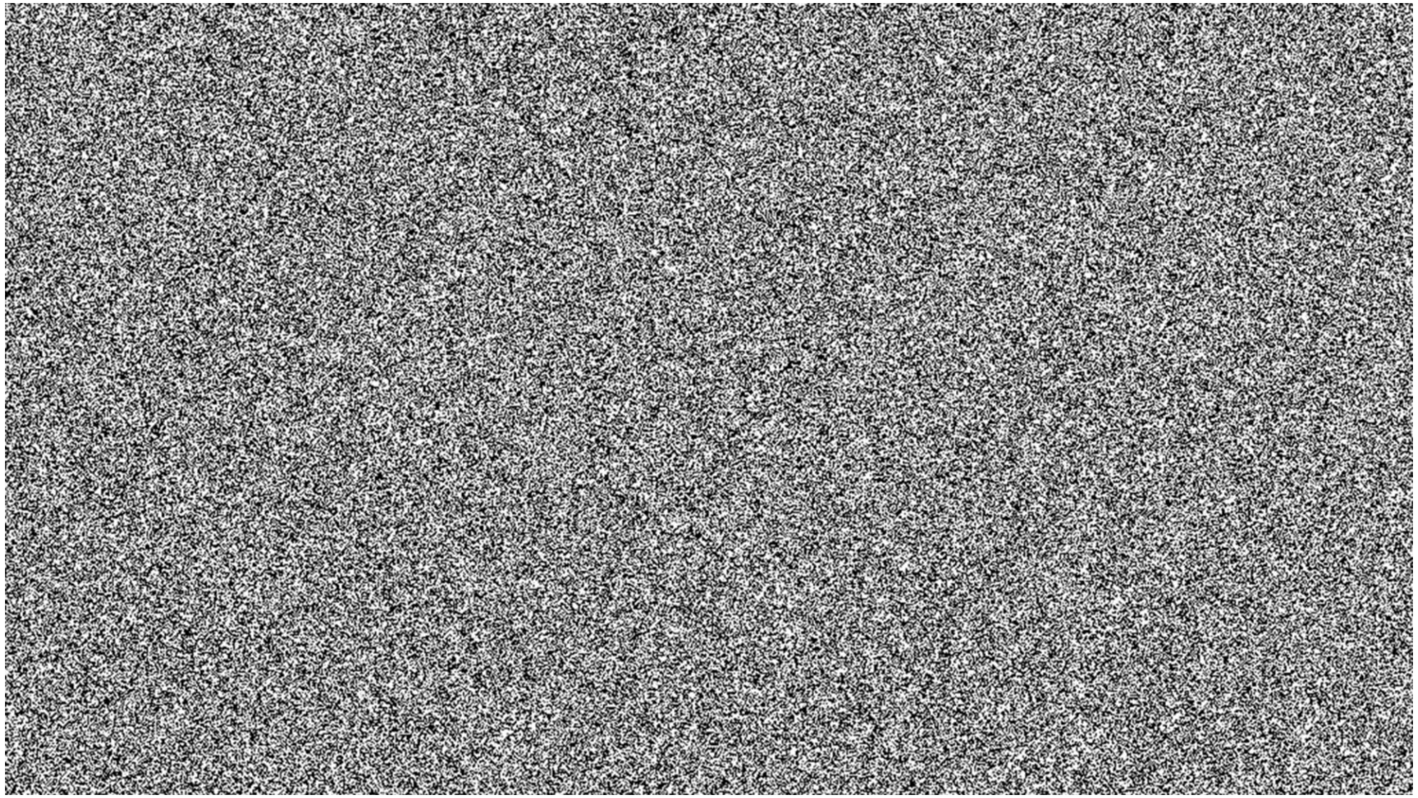
$p_0 = 1$ et $p_i = 0, \forall i \neq 0$

$H(X) = -\log_2(1) = 0$

$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

Calcul de l'entropie d'une image

- Quelle est l'entropie d'une image couverte de bruit blanc ? (distribution uniforme)

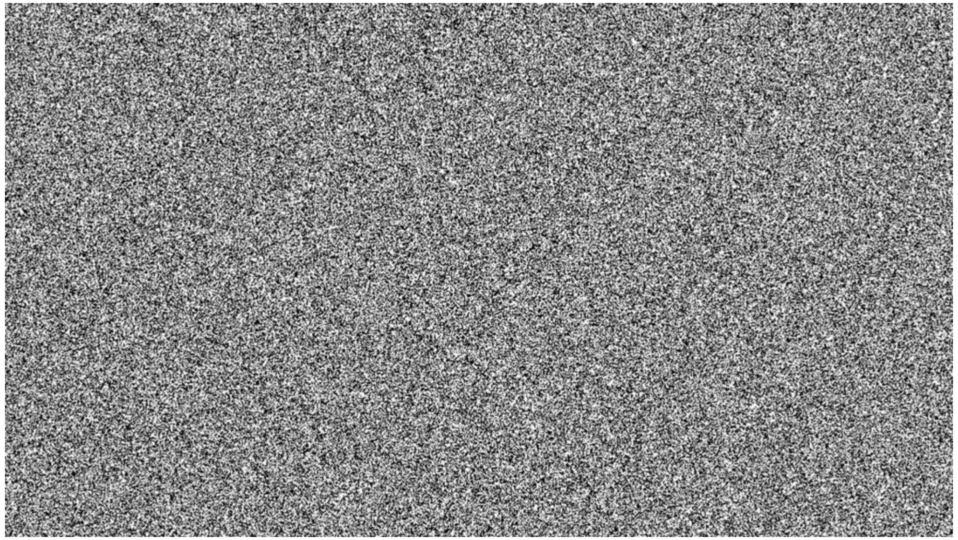


Calcul de l’entropie d’une image

- Quelle est l’entropie d’une image couverte de bruit blanc ?

$$p_i = \frac{1}{256}, \forall i \in [0,255]$$

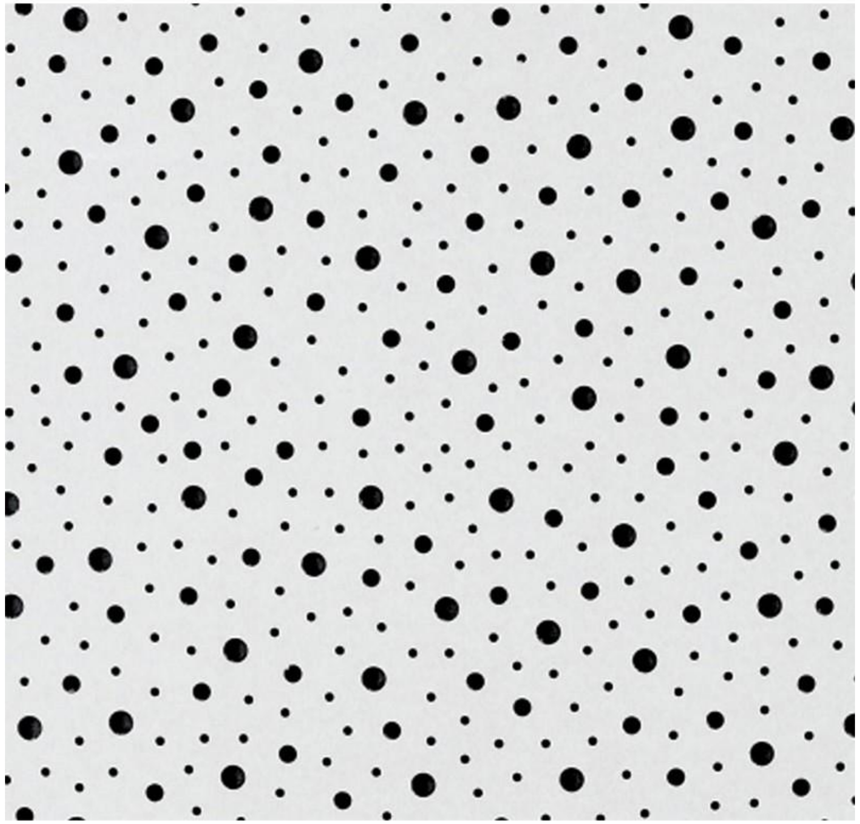
$$H(X) = -\log_2\left(\frac{1}{256}\right) = 8$$



$$H(I) = -\sum_{i \in I} p_i \log_2(p_i)$$

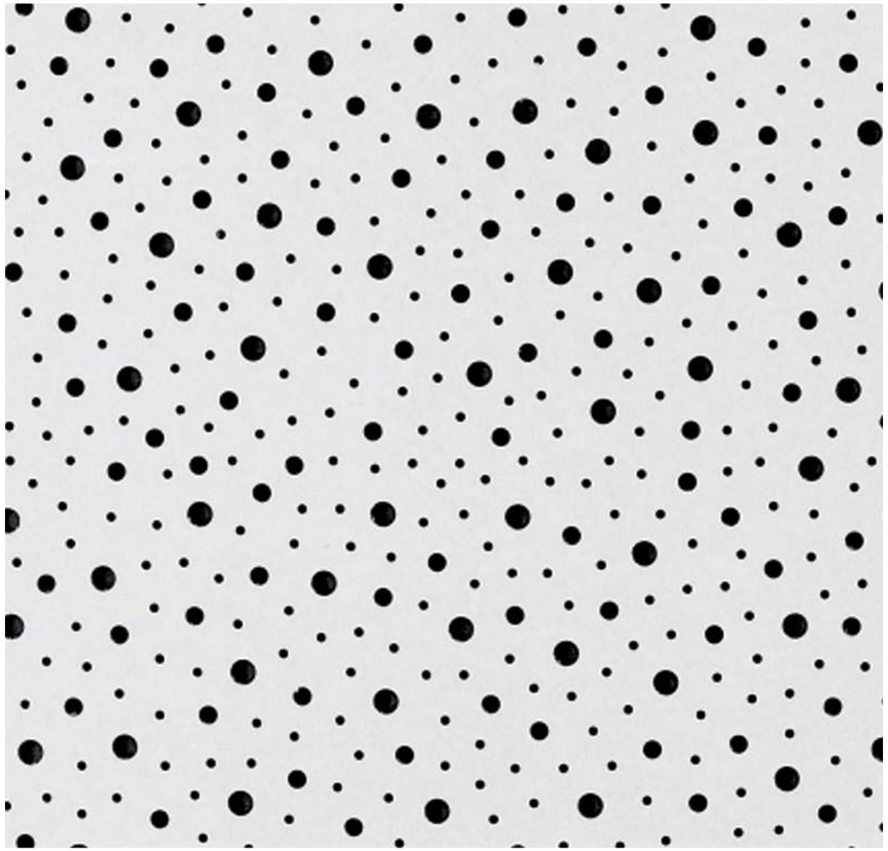
Calcul de l’entropie d’une image

- Quelle est l’entropie d’une image contenant 80% de pixels blancs et 20% de pixels noirs ?



Calcul de l'entropie d'une image

- Quelle est l'entropie d'une image contenant 80% de pixels blancs et 20% de pixels noirs ?



$p_0 = 0.2$ et $p_{255} = 0.8$

$p_i = 0, \forall i \neq \{0{,}255\}$

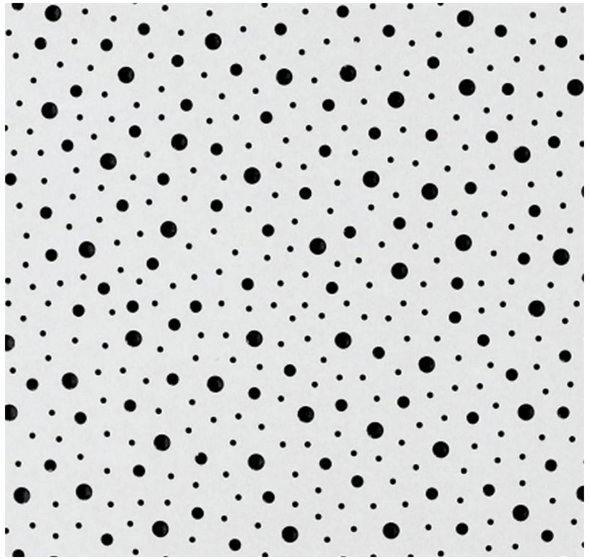
$H(X) = -0.2\log_2(0.2) - 0.8\log_2(0.8)$

$H(X) \approx 0.7219$

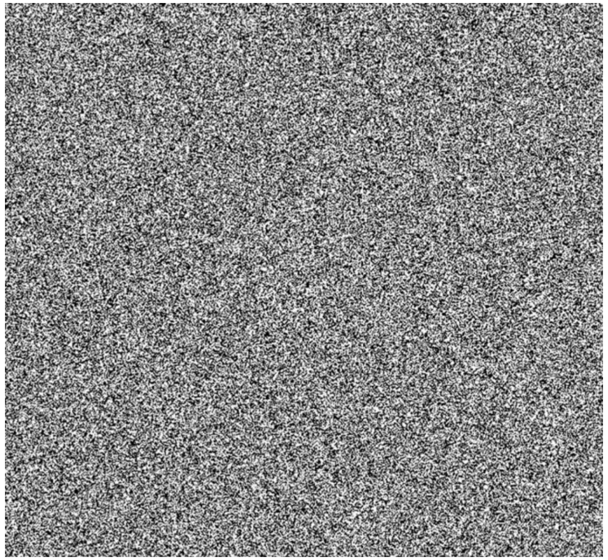
Calcul de l’entropie d’une image

Récapitulatif :

$H(X) = 0$



$H(X) \approx 0.7219$

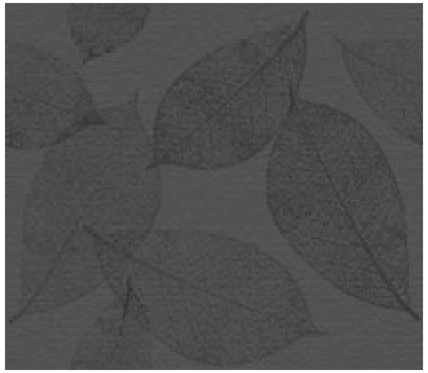


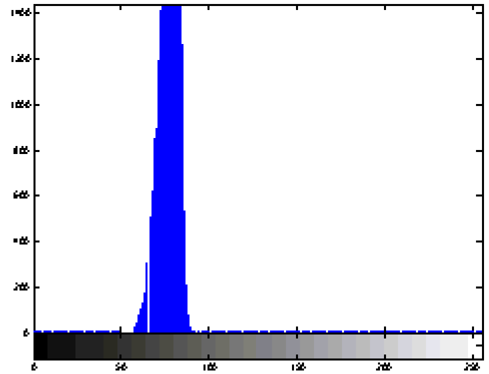
$H(X) = \log_2(256) = 8$

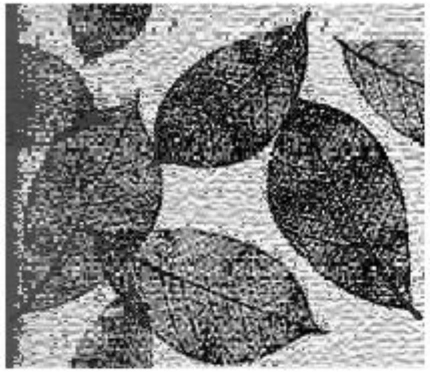
Pour les images aussi : $H(X) \leq \log_2 m$

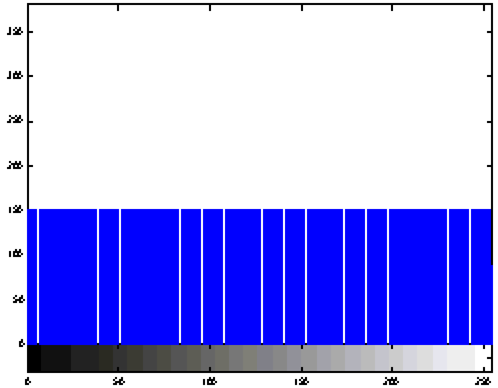
Calcul de l’entropie d’une image

Voici des images et leur histogramme :





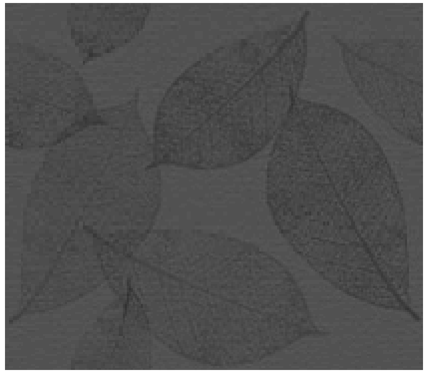


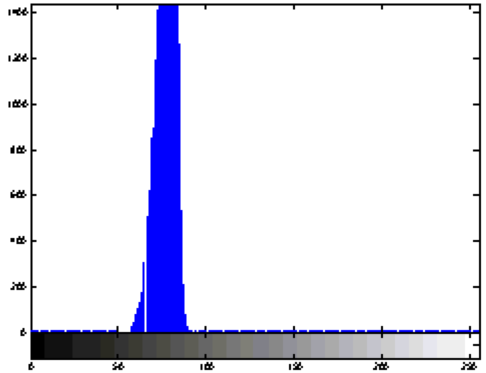


Quelle image a l’entropie la plus haute ?

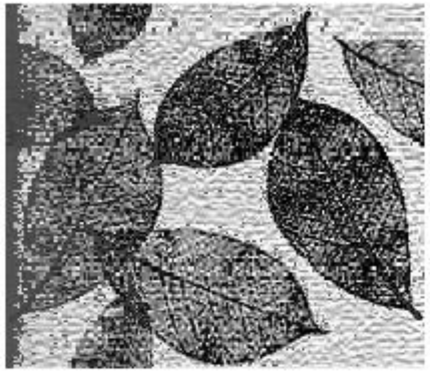
Calcul de l’entropie d’une image

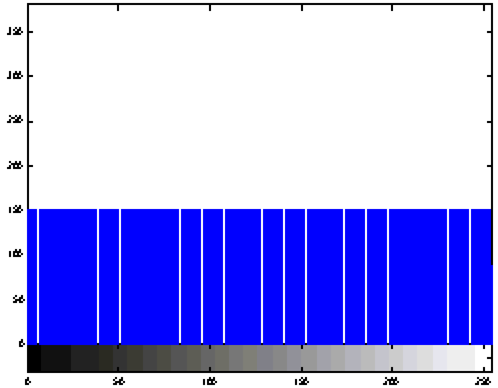
Voici des images et leur histogramme :





$H(X) = 3.2423$





$H(X) = 5.548$

L’histogramme le plus plat correspond à l’entropie la plus grande, car il est le plus proche de la variable aléatoire uniforme (borne sup de H).

Il y a une incertitude maximale sur la valeur qu’aura un pixel pris au hasard.

Entropie

Résumé :

- $H(I) = -\sum_{i \in I} p_i \log_2(p_i)$
- L'entropie mesure la quantité moyenne d'information d'un ensemble d'évènements, et son incertitude.
  - Entropie nulle : aucune incertitude, on peut deviner la valeur d'un tirage de la variable aléatoire
  - Entropie maximale : incertitude maximale, on ne peut pas deviner la valeur d'un tirage de la variable aléatoire

Information partagée par deux variables