

Information, Calcul et Communication

Théorie: Représentation de l'Information (1)

R. Boulic

Existe-t-il une représentation universelle de l'information ?

Plan

Lien avec les Leçons précédentes

- **Rappel des domaines d'applications**
- **Une représentation est une convention**
- **Vers l'unité élémentaire d'information (exercices)**

Manipulation sur les nombres entiers

- **Opérations et domaine couvert**

La virgule flottante: Pourquoi ? Comment ?

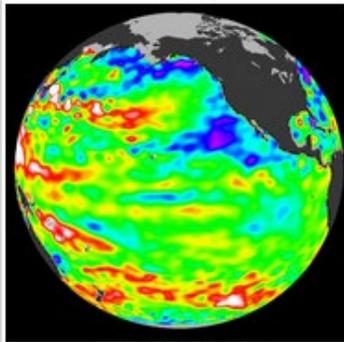
- **Un exemple qui pose problème**

Retour à la représentation des symboles

- **De l'alphabet aux idéogrammes**

Lien avec l'introduction

Domaines d'application



En 2003, mise en évidence du "Niño" par simulation numérique de la circulation océanique - © INRIA / Projet IDOPT

Calcul scientifique /Simulation

-> *nombres*

Conduite de processus

-> *signaux (mesures, contrôle...)*

Gestion d'information

-> *texte*



Google datacenter



Honda Asimo

Une représentation est une convention



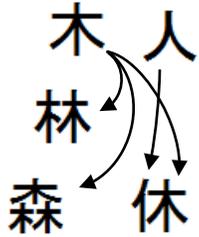
- pour faciliter l'activité d'un groupe d'utilisateurs
- correspondance entre un ensemble de signes et leur signification.

Fragilité de cette convention: langues mortes, codes perdus, etc...

Il n'existe pas de représentation universelle

- standard *de facto* = porté par le marché, l'usage (ex: pdf)
- standard *de jure* = normalisation (IEEE, ACM, ISO...).

exemples: alphabet romain, chiffres indo-arabes, code de la route, papier monnaie



A B C ...

0 1 2 3...

Vers l'unité élémentaire d'information

214 motifs graphiques, appelés des clefs, ont été utilisés pour construire ~100.000 idéogrammes chinois

Les **26 lettres** de l'alphabet latin ont été utilisés pour construire ~1.000.000 mots des langues occidentales

Les **10 chiffres** indo-arabes permettent de construire une infinité de nombres (et même d'encrypter tous les mots!)

Question: quel est le système de signes le plus simple permettant de conserver la même efficacité d'expression que les 10 chiffres ?

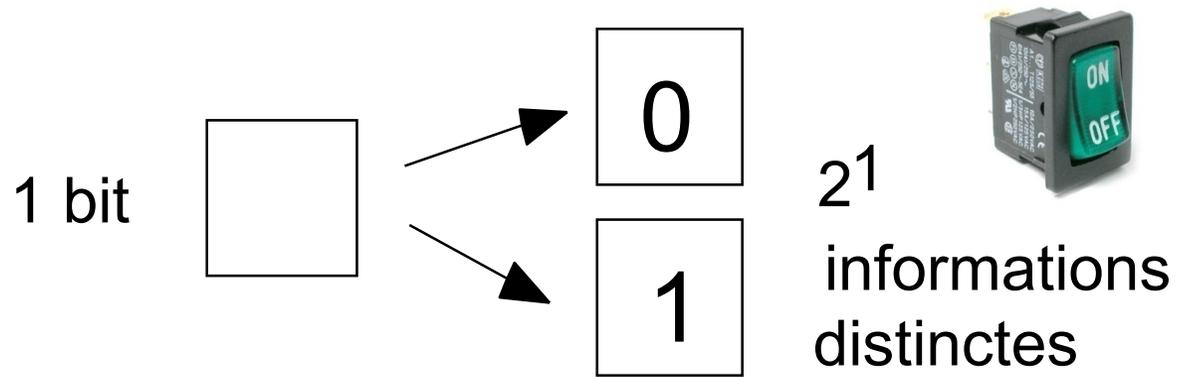
Réponse: un système de **2 symboles**

Toute information peut être représentée à l'aide d'un ensemble **d'éléments binaires**

Par convention, un **élément binaire** vaut **0** ou **1**.

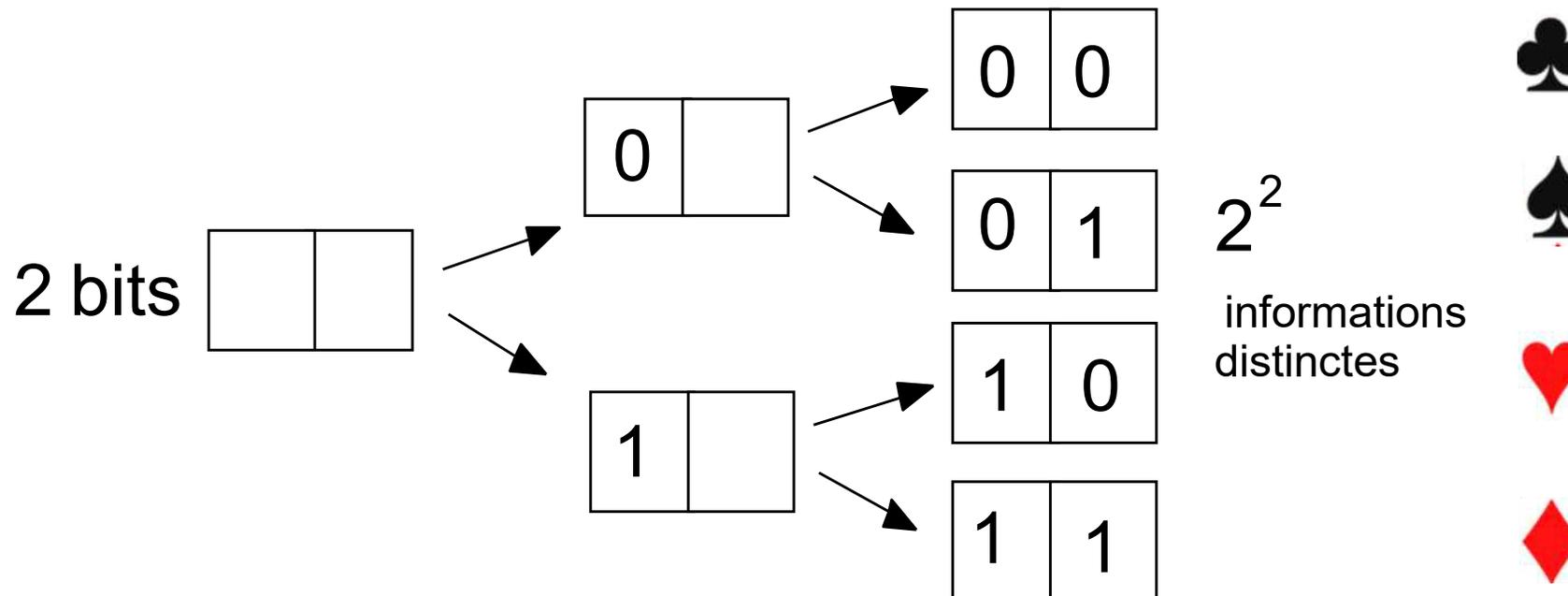
Par la suite on utilise l'expression anglaise "**binary digit**" ou **bit** en abrégé

Dans cette leçon nous faisons abstraction de la manière dont les éléments binaires sont réalisés (états magnétiques, tensions, courants, etc..). Cela sera abordé dans le Module3.



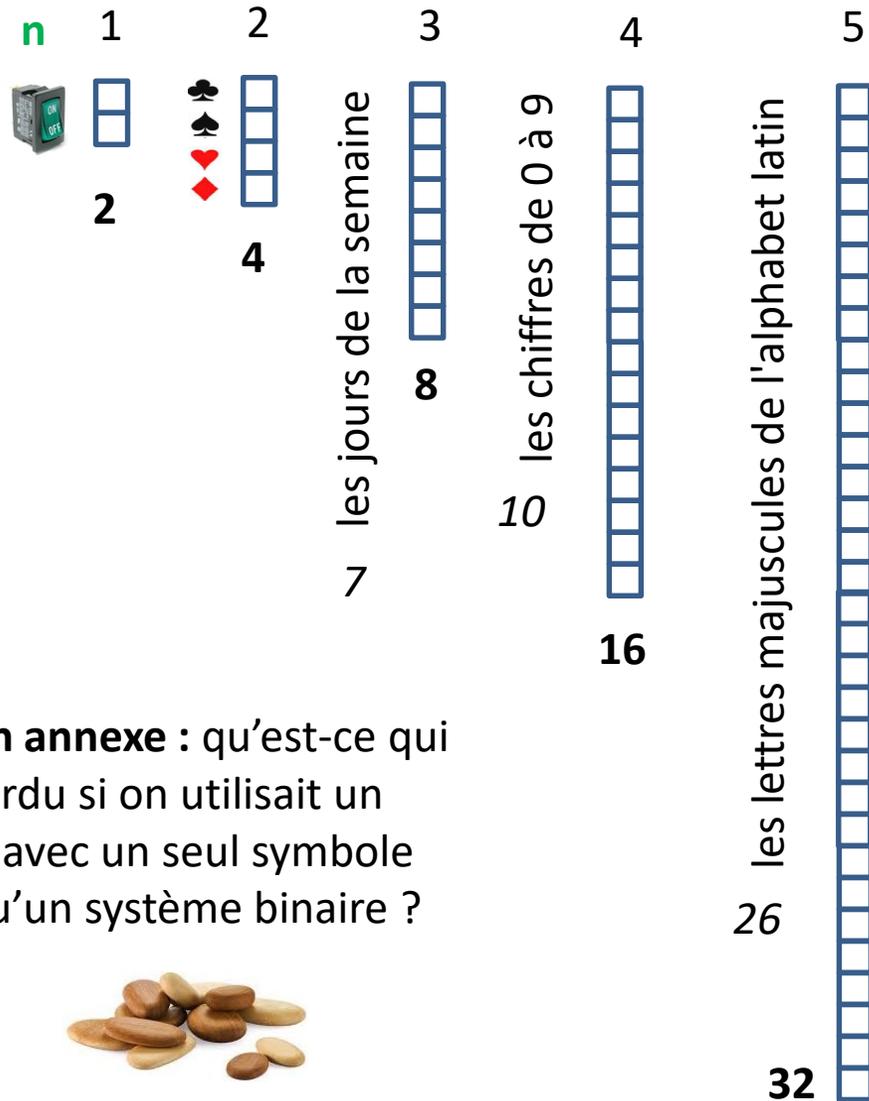
Capteurs à seuil

Comment représenter plus d'informations ?



n bits permettent de construire 2^n **combinaisons** distinctes pour représenter 2^n informations distinctes

Réciproquement, 2^n informations distinctes sont représentables par $\log_2(2^n) = n \log_2(2) = n$ bits



Exercice:

Combien de bits suffisent pour représenter :

- les jours de la semaine :
- les chiffres de 0 à 9 :
- les lettres de l'alphabet:
 - Majuscules
 - Minuscules + Majuscules
 - Min + Maj + chiffres + signes ...

règle générale:

Pour **K informations distinctes**, le **nombre de bits n** suffisant pour représenter ces informations est l'entier supérieur ou égal à **$\log_2 K$**

Question annexe : qu'est-ce qui serait perdu si on utilisait un système avec un seul symbole plutôt qu'un système binaire ?



n bits permettent de représenter **2^n** informations distinctes

n	2^n
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
10	1024
20	1048576
30	1073741824
32	4294967296

Bonne pratique pour estimation rapide:

$$2^{10} = \text{kibi (Ki)} \approx 10^3 = \text{kilo (k)}$$

$$2^{20} = \text{mébi (Mi)} \approx 10^6 = \text{méga (M)}$$

$$2^{30} = \text{gibi (Gi)} \approx 10^9 = \text{giga (G)}$$

$$2^{32} = 2^{30+2} = 2^{30} \cdot 2^2 \approx 4 \text{ G}$$

Organisation de l'information

Convention: on appelle **byte** un groupe de 8 bits (**octet**).

L'octet est la **brique de base de la mémoire centrale**

Les représentations les plus courantes pour l'information exploitent l'octet comme élément de base.



Résumé de l'introduction

Existe-t-il une représentation universelle de l'information ?

Une représentation est une **convention humaine** d'interprétation d'un ensemble de signes. Sa force est directement liée au nombre de personnes qui la partage, d'où l'importance des **standards** (ex: code ASCII, UTF).

Par quels moyens peut on représenter des symboles et des nombres ?

La représentation binaire suffit pour représenter un nombre arbitrairement grand de signes. Par convention nous utilisons les symboles 0 et 1.

D'un point de vue technique, le système binaire présente également des avantages pour réaliser des solutions robustes à toutes sortes de perturbations.