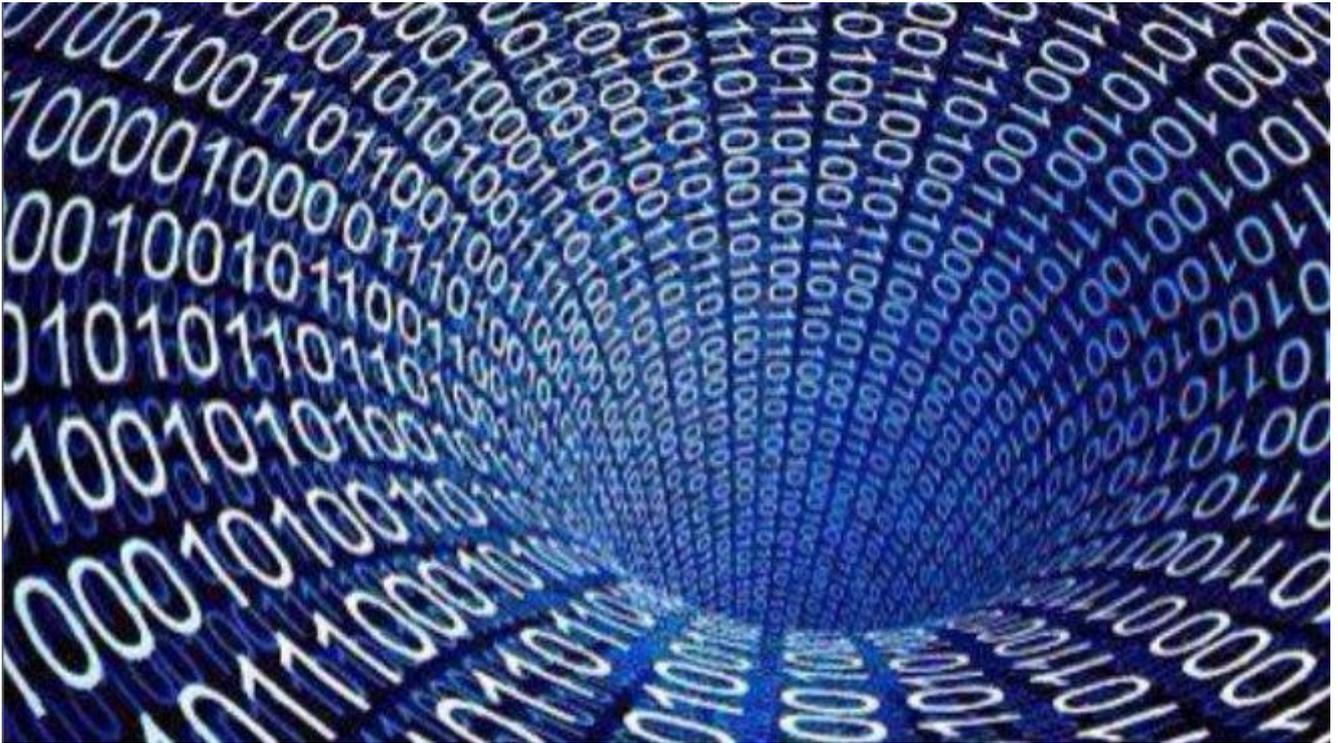


LE LANGAGE BINAIRE



I.	<i>Un peu de Théorie de l'Information.</i>	2
II.	<i>Codage de l'Information.</i>	4
III.	<i>Le Langage Binaire.</i>	6
IV.	<i>L'autre unité de quantité d'Information : l'Octet.</i>	7
V.	<i>Bilan des acquisitions.</i>	10

- Matériel :
- Pré-requis pour prendre un bon départ :

Différence entre les chiffres et les nombres.				
Puissances de 2.				

« L'Informatique est un domaine d'activité scientifique, technique et industriel concernant le traitement automatique de l'Information par l'exécution de programmes informatiques par des machines. » (Wikipédia)

Nous avons déjà vu le traitement automatique (partie Algorithmique) et l'exécution de programmes (partie Langages et Programmation). Intéressons-nous plus en détail à l'Information.

I. UN PEU DE THEORIE DE L'INFORMATION.

A. L'Information, c'est quoi ?

C'est une question très difficile et la littérature est abondante sur ce sujet, donc je ne me hasarderai pas à en donner une définition catégorique. Essayons quand même :

« L'Information est contenue dans toutes choses matérielles ou immatérielles (situations, évènements) : c'est tout ce qui permet de caractériser cette chose ou cet évènement. »

Plusieurs remarques se dégagent :

- L'Information est immatérielle : ce ne peut être une masse ou une énergie.
- L'Information n'est pas la chose mais elle lui donne ses caractéristiques, son sens, son essence ?
- L'Information n'a d'intérêt que parce qu'elle peut être communiquée.

B. Représentations de l'Information :

Intéressons-nous à la dernière remarque ci-dessus.

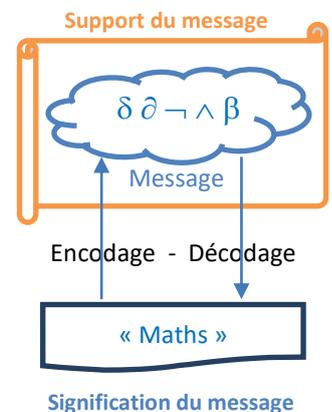
➤ Pour pouvoir être communiquée, cette Information pure doit être inscrite dans le monde réel en étant associée à un objet concret appelé **support**. Ce support va subir une modification physique.

C'est cette modification physique du support qui va représenter l'Information.

Exemples : dans une représentation textuelle de l'Information, le support sera le papier et les modifications physiques seront les traces d'encre sur ce papier ; dans une représentation chantée de la même Information, le support est l'air et les modifications physiques sont les vibrations sonores dues aux mots et à la musique.

➤ Bien sûr, il faut que ces modifications physiques (traits de stylo, mots prononcés) obéissent à des règles pour que ces représentations de l'Information conservent du sens : on parle de **codage de l'Information**.

- **Toute représentation de l'Information est l'association d'un support et d'un codage.**
- **Les modifications physiques du support forment des signes ou signaux qui constituent un message.**
- **L'Information pure est la signification de ce message.**
- **Le codage est l'ensemble des opérations (encodage-décodage) faisant correspondre les signes constituant le message avec la signification de ce message.**



Les ordinateurs manipulent les signes des représentations sans comprendre le sens du message. C'est là leur force : ils traitent les signes à toute vitesse sans être arrêtés par les significations, sans états d'âme !

C. Problématique des représentations :

Souvent, on confond l'Information (le contenu informatif de la chose) et les représentations qu'on peut en faire afin de communiquer cette Information. Par exemple : l'évènement « Seconde Guerre Mondiale » contient une grande quantité d'informations. Un film, un livre ou un site internet ne peuvent en être que des représentations partielles.

1. Il est impossible de représenter parfaitement toute l'Information :

Une représentation est forcément partielle, subjective et imparfaite :

- partielle : l'Information contenue étant quasi infinie, celui qui veut la communiquer doit opérer des choix restrictifs dans sa représentation.
- subjective : qui dit choix dit forcément subjectivité de ce choix. Une représentation ne peut donc être totalement objective malgré toute l'honnêteté qu'on peut y mettre.
- imparfaite : même la partie retranscrite ne peut être parfaite ! En effet, la transcription de l'Information sur un support selon un certain codage fait plus ou moins perdre de l'Information, donc de la signification.

Exemples : Un tableau est une représentation en 2D d'une réalité en 3D ;
.....

2. Il n'y a pas unicité des représentations :

- Variété des supports (disques durs, livres,).
- D'un côté, cette variété produit une richesse tant patrimoniale que culturelle ou pédagogique. Mais d'un autre côté, cela oblige les citoyens à maîtriser les langages liés à ces supports.
- Variété des codages pour un même support :

3. Quelle problématique se dégage concernant les représentations ?

➤ Nous laisserons à d'autres champs d'étude les problèmes d'ordre philosophique (partialité et subjectivité des représentations) et éducationnel (maîtrise des différentes formes de communication liées aux différents supports) pour nous concentrer sur le problème technique du codage de l'Information sur un support.

- Une foule de questions se posent alors :
 - S'il existe plusieurs codages de l'Information pour un même support, par quels critères les différencier ?
 - Peut-on décider qu'un codage est meilleur qu'un autre et selon quel critère ?
 - L'Information contient forcément des données. Quels types de données sont représentables ?
 - Existe-t-il un codage qui peut « unifier » des types de données a priori différents ?

II. CODAGE DE L'INFORMATION.

A. Choix du codage :

Un codage fait la correspondance entre les signes ou signaux constituant le message et le sens de ce message.

➤ Les critères vous semblent importants pour le choix d'un bon codage ?

.....
➤ Parmi les 9 codages suivants quel est celui qui vous semble le mieux correspondre à ces critères :

- | | | |
|----------------------------------|-----------------------------------|--------------------------------------|
| codage à base de signes inventés | codage à base d'odeurs | codages à base de phrases |
| codage à base de mots | codage à base de sons | codages à base de mimiques du visage |
| codage à base de nombres | codage à base de signaux lumineux | codage à bases A, T, C, G. |

Donc, du fait de leur universalité, de leur caractère simple et pratique, de leur non ambiguïté, de leurs propriétés structurelles et calculatoires, on choisit évidemment un codage à base de

B. Types de données représentables avec un codage numérique :

- Toutes les données de type numérique évidemment !
- Les textes ?
- Les sons ?
- Les images ? Les couleurs ?
- Les odeurs ?
- Les sensations ? Les sentiments ? L'Intelligence ?

Le Monde est-il nombre ? On peut se poser la question ...

C. Quel codage numérique ?

On va donc utiliser des nombres pour transcrire la foultitude d'informations.

Mais comment ces nombres eux même seront-ils écrits ? Se pose donc la question de combien de on va utiliser pour écrire ces nombres.

1. Nombre minimum de chiffres nécessaires au codage numérique :

- Il faut évidemment au moins un chiffre pour écrire des nombres. Un chiffre est-il suffisant ?
- Supposons donc qu'il n'y ait qu'un seul chiffre pour coder l'Information, le chiffre 1 par exemple.

Par exemple une vache serait codée 1, une poule : 11, une image : 111, une autre image :

Comme il n'y a qu'un seul chiffre, la seule façon de différencier les choses serait de jouer sur la du nombre.

➤ Donc toute chose serait représentée par une suite plus ou moins longue de chiffres 1 !

Quel problème cela pose-t-il concrètement ?

Donc le nombre minimum de chiffres pour coder l'Information n'est pas un mais au minimum.

➤ Mais peut-être ce nombre minimum de chiffres est 3 ou 4 ou 10 ou plus !! Voyons cela dans ce qui suit.

2. Pas 1, pas 10 mais :

➤ L'idée la plus naturelle serait de choisir chiffres pour le codage numérique de l'Information, choix qui est la base de notre numération.

Mais est-ce un bon choix ? Dans un monde abstrait et parfait comme celui des Maths : ! Mais qu'en est-il dans le monde concret et réel (imparfait) des machines ?

➤ Les ordinateurs et par extension toute machine dédiée au traitement de l'Information (systèmes embarqués, robots, smartphones, automates etc.) fonctionnent à l'électricité.

Donc chaque chiffre permettant de coder l'Information va être représenté par un état électrique, la tension électrique en l'occurrence.

➤ Si on choisit 10 chiffres pour coder l'Information, alors il faut qu'il y ait 10 tensions significatives différentes qui leur correspondent. Plus exactement, vu que la tension est continue : les 10 chiffres doivent correspondre chacun à 10 intervalles significatif de tensions.

Par exemple la plage [0 volt ; 1 volt[correspondrait au chiffre 1, la plage [1,1 volts ; 3 volts[correspondrait au chiffre 2, etc.

Et donc, si l'ordinateur « mesure » 0,95 volts, il « lit » théoriquement le chiffre

➤ Seulement voilà, nous sommes dans le monde réel ! Et dans le monde réel imparfait, qui dit « mesure », dit imprécisions liées aux parasites électriques, à la précision du mesureur de la machine etc.

Supposons que cette imprécision soit de $\pm 0,2$ volts. En reprenant l'exemple précédent, l'ordinateur « lit » toujours le chiffre 1, mais peut-être était-ce en réalité le chiffre que l'ordinateur devait lire !

Il y a donc ambiguïté de la lecture ! Impardonnable, cela entraînerait bug sur bug dans les programmes !

➤ Il n'y a que 2 façons possibles de remédier à ce problème d'imprécision de la mesure :

- soit diminuer l'imprécision liée à la mesure de la tension :

Impossible car la bonne lecture dépendrait alors des mesureurs de chaque machine !

- soit augmenter la distance entre les intervalles significatifs de tension.

Ce qui revient forcément à diminuer le nombre de chiffres codant l'Information !

➤ La distance entre les intervalles significatifs de tension est maximum lorsqu'il n'y a que chiffres codant l'Information !

Ces 2 seuls chiffres codant numériquement l'Information seront donc :

- le chiffre qui correspondra à la tension 0 (pas de courant).
- le chiffre qui correspondra à la tension d'alimentation (il y a du courant).

Voilà donc pourquoi les ordinateurs fonctionnent en binaire.

Ce système de retranscription et de transmission de l'Information basé sur les chiffres 1 et 0 constitue un langage : le Langage

III. LE LANGAGE BINAIRE.

A. Définitions et vocabulaire :

❶ Le langage ayant pour base les deux chiffres et s'appelle le **Langage**

Ce langage binaire est le langage qui a été choisi (voir II.A p.4) pour les machines : c'est pourquoi on parle aussi de **Langage Machine**.

❷ En langage binaire, toute information est représentée par **une suite de 0 et/ou de 1**.

❸ Cette suite de 0 et/ou de 1 s'appelle un **mot binaire**.

Ces mots binaires étant constitués de 0 et de 1 peuvent être aussi vus comme des nombres. C'est pourquoi on ne dit presque jamais mot binaire mais **nombre binaire**.

Là réside le deuxième grand avantage du langage binaire (outre la simplicité de ses représentations) : le langage binaire bénéficie des avantages structurels liés aux nombres : en particulier les opérations.

B. Le bit :

❹ Chaque chiffre d'un mot / nombre binaire s'appelle un **bit** (**binary digit**, chiffre binaire en anglais).
Le bit est donc la plus petite quantité d'information manipulable par une machine.

Le bit est l'Unité élémentaire de quantité d'Information.

Le bit ne peut prendre que deux valeurs : ou

❺ Le bit commençant un nombre binaire (bit le plus à gauche) s'appelle le **bit de poids le plus fort**.

❻ Le bit terminant un nombre binaire (bit le plus à droite) s'appelle le **bit de poids le plus faible**.

Cette dualité du bit peut correspondre à plusieurs types de situations :
 Numérique (0 ou 1) ; Logique (faux ou vrai) ; Interrupteur (ouvert ou fermé) ; Magnétique (nord ou sud) ;
 Optique (absence ou présence de lumière) ; Cartes perforées ou CD (absence ou présence de trous) etc.

C. Nombre maximum d'états distincts codables avec N bits :

La mémoire des ordinateurs est constituée d'une multitude de petits circuits électroniques ne pouvant être chacun que dans deux états représentés par 0 (absence de tension électrique) ou 1 (présence d'une tension électrique). **Un tel circuit à deux états s'appelle *circuit mémoire un bit* ou *commutateur*.**

L'assemblage de ces commutateurs permet de réaliser les opérateurs logiques de base (opérateurs ET, OU, NON) qui permettront à leur tour d'assembler des systèmes complexes (additionneurs etc.).

Soit donc une mémoire d'ordinateur composé de plusieurs commutateurs (circuits mémoire un bit).

L'état de cette mémoire se traduit par une suite finie de 0 et de 1. Par exemple, le mot binaire 100 décrit l'état d'une mémoire composé de trois commutateurs, respectivement dans les états 1, 0 et 0.

❶ Un circuit mémoire est constitué de 6 commutateurs, un autre de 10 commutateurs.
 Quel est le nombre d'états possibles de la mémoire dans chaque cas ?

❷ Plus généralement pour un ordinateur dont la mémoire est constituée de N commutateurs.

D. Nombre de bits nécessaires au codage de k valeurs distinctes :

① La Nature a choisi une autre base pour le codage de l'ADN. Laquelle ?

Si l'on devait recoder en langage machine ce codage en base 4, combien de bits faudrait-il ?

② Un circuit mémoire doit être capable de manipuler les représentations binaires d'un jeu de 52 cartes.

De combien de commutateurs minimum doit être composée cette mémoire ?

Plus généralement, calculer le nombre minimum N de bits pour coder k états distincts ?

③ Les performances du tableur Excel de Microsoft ont augmenté avec la montée en gamme des processeurs.

Voici quelques spécifications pour Excel 2003 et Excel 2016 :

	Excel 2003		Excel 2016	
	spécification	codée sur ... bits	spécification	codée sur ... bits
Nb de lignes	65 536 = 2 ¹⁶		1 048 576	
Nb de colonnes	256		16 384	
Nb max de caractères par cellule	32 767		32 767	
Nb max de caractères par formule	1 024		8 192	

Le codage en binaire explique souvent pourquoi on n'a pas de nombres « ronds » comme spécifications des programmes.

IV. L'AUTRE UNITE DE QUANTITE D'INFORMATION : L'OCTET.

A. Regroupement des bits par paquets :

➤ On rappelle qu'un mot binaire est une suite plus ou moins longue de et de

Voici un motif binaire (c-à-d un mot binaire, un nombre en langage machine) : 1000111111000000

Sur combien de bits est-il codé ? Dit autrement combien de chiffres binaires a-t-il ?

Est-il facile à lire et à écrire ?

➤ De même que pour écrire plus lisiblement 123456789 en écriture décimale, on écrit 123 456 789, on va regrouper par paquets les bits. Mais par paquets de combien ?

La réponse à cette question est liée intimement à l'histoire des processeurs des ordinateurs. Comme le codage binaire est un système à base 2, les processeurs (et les systèmes d'exploitation associés) ont traité d'abord des paquets de 1 bit, puis 2 bits, puis 4 bits, puis bits, puis bits etc.

➤ Historiquement, le choix a été fait d'écrire les nombres binaires par paquets de bits.

Un regroupement de 8 bits s'appelle un Octet.

L'écriture en octets rend l'écriture binaire plus lisible.

➤ Application :

1. Réécrire le nombre précédent en faisant apparaître les octets :
2. Réécrire le mot binaire 1110001011 sur 2 octets quitte à compléter la nouvelle écriture par des 0 supplémentaires à gauche :
3. Les systèmes modernes d'exploitation (Windows 10 par exemple) manipulent des mots binaires codés sur 64 bits. Combien cela fait-il d'octets ? octets.

B. Définition de l'Octet :

Rappel : d'après la p.6, la plus petite Unité de quantité d'Information est le

L'Octet, outre son caractère pratique, permet de définir une nouvelle unité de quantité d'Information :

Définition : L'Octet est l'autre unité de quantité d'Information : celle formée à partir de 8 bits.

L'Octet est l'unité qui a été choisie pour représenter les capacités mémoire.

Attention : En anglais un octet s'écrit *byte* (parfois *Byte* avec une Majuscule), ce qui ressemble à bit !

C. Multiples de l'Octet :

➤ Depuis Décembre 1998, l'International Electrotechnical Commission (IEC) a grandement simplifié les conversions entre les multiples de l'Octet : celles-ci doivent suivre les mêmes principes que les conversions en système décimal (multiplication ou division par 10 pour passer d'une colonne à une colonne adjacente).

Giga octet (Go)			Méga octet (Mo)			kilo octet (ko)			octet

D'après ce tableau de conversion (attention, Majuscules sauf pour kilo-octet) :

1 ko = 1 octet = 10³ octets = mille octets.

1 Mo = 1 octet = 10⁶ octets =

1 Go = 1 octet = 10⁹ octets =

1 To = 1 octet = 10¹² octets =

➤ Attention !

Avant cette standardisation simplificatrice de 1998, 1 kilo octet valait 1 024 octets, 1 Mo valait 1 024 ko, 1 Go valait 1 024 Mo, 1 To valait 1 024 Go etc.

V. BILAN DES ACQUISITIONS.

➤ A la fin de ce livret, je dois savoir :

Représentation et codage en binaire (7 compétences)				
Expliquer pourquoi les machines travaillent en binaire.				
La définition d'un bit, d'un octet, d'un mot binaire.				
Trouver combien de bits il faut pour représenter n états distincts.				
Calculer en octets, ko, Mo, Go, To etc. le poids d'une quantité d'informations.				

➤ Nous avons maintenant toutes les armes nécessaires pour commencer à numériser nos données.

Evidemment, comme d'habitude, on commencera par le type de données le plus simple : les

.....

Suite donc au prochain livret : « Numérisation des ».

➤ Perles des hotlines informatiques :

« Secrétaire - J'ai un problème avec Windows.

Hotline - Qu'avez-vous sur l'écran ?

Secrétaire - Euh... un pot de fleur.

Hotline - Non, je veux dire "Qu'est-ce qui est écrit sur l'écran ?".

Secrétaire - Ha d'accord... euh... Sony. »

« Client - Vous me dites "pas de majuscules pour le mot de passe", c'est bien ça ?

Hotline - Exact.

Client - Et, les chiffres, je les mets en minuscule aussi ? »