

Les fichiers informatiques

« *Le langage d'un outil informatique ne comprend que 2 lettres (0 et 1) »* »

Pour tout exprimer ou **enregistrer**

Il était une fois une petite fille de Village, la plus jolie qu'on eût su voir ;

texte



Images



films



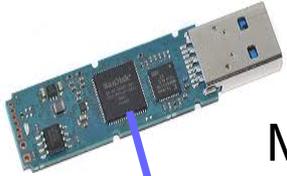
musique

On parle de **langage binaire**



exemple avec le « principe » du stockage des fichiers

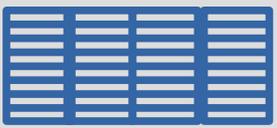
Périphérique de stockage



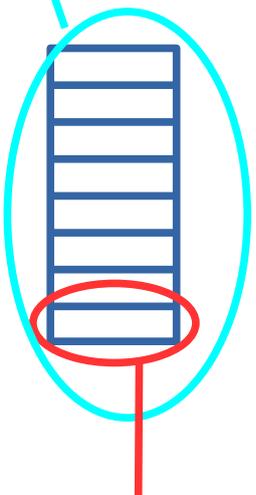
disque dur
Clef USB
Mémoire vive

Clef 4Go : *32000000000* de « cases »

Clef USB



Des milliards de « cases » Appelé : BIT



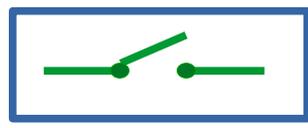
Chaque case

Comme un interrupteur



1

- Soit fermé :
« courant électrique » ~1



0

Soit ouvert : « pas de courant électrique » ~ 0

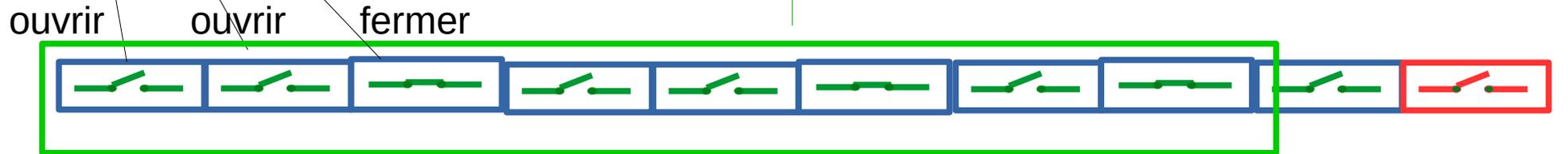
Enregistrer un fichier c'est comme ouvrir ou fermer des interrupteurs des différents « cases » qui composent l'unité de stockage.

Cf : Petite anatomie d'unités de stockage

Par exemple un fichier informatique enregistré dans une clef USB

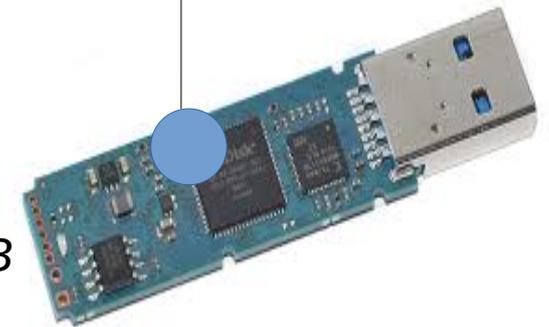
001001010100101011100011111...

1 « case » = 1 Bit : 0 ou 1



X 1000000000

En fait, il n'y a pas vraiment d'interrupteurs, comme ceux que l'on voit sur les murs.. il s'agit d'une image pour comprendre le principe d'enregistrement des fichiers...



Cf : Petite anatomie d'unités de stockage

Capacité clef USB

4 Go = = 32000000000 cases

Plus petite unité d'information

binary Unit

Chiffre binaire

« Une lettre de l'ordinateur » : **1 Bit**

Une information numérique de 16 Bit

0110101011000101



exemple : Texte (simple)

« *Le langage d'un outil informatique ne comprend que 2 lettres (0 et 1)* »

Cf : processeur

Pour tout exprimer ou enregistrer

Exemple du texte

*Il était une fois une
petite fille de Village, la
plus jolie qu'on eût su
voir ;*

Comment ?

Questions préalables

Lorsque vous lisez un texte
quels sont les symboles
différents que vous utilisez ?

Combien y en a t-il (environ) ?



Réponse

Beaucoup

a,b,c,d,.....z

+ *A,B,C,D....Z*

+ *é,à,è: ù*

+ *0,1,2,3...9*

+ *<, >, +, =*

+ *« ; ! § \$...*

Entre 200 et 250



8 bits : 256 symboles différents

Certains peuvent apparaître encore



Apparu en 1998

Activité de réflexion

1) Comment coder (en binaire)
n'importe quel texte de manière
universelle pour que tous les
ordinateurs puissent l'interpréter ?

Démarche d'investigation

Question
Problème



Hypothèses
Conjoncture
Idée



Évaluation
acquis

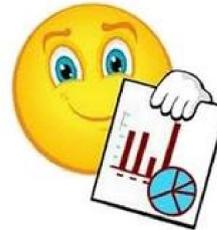


Synthèse
formalisation



Expérimentation
réalisation matérielle,
Observation directe
recherche sur documents
une enquête, une visite

Confrontation



Conclusion



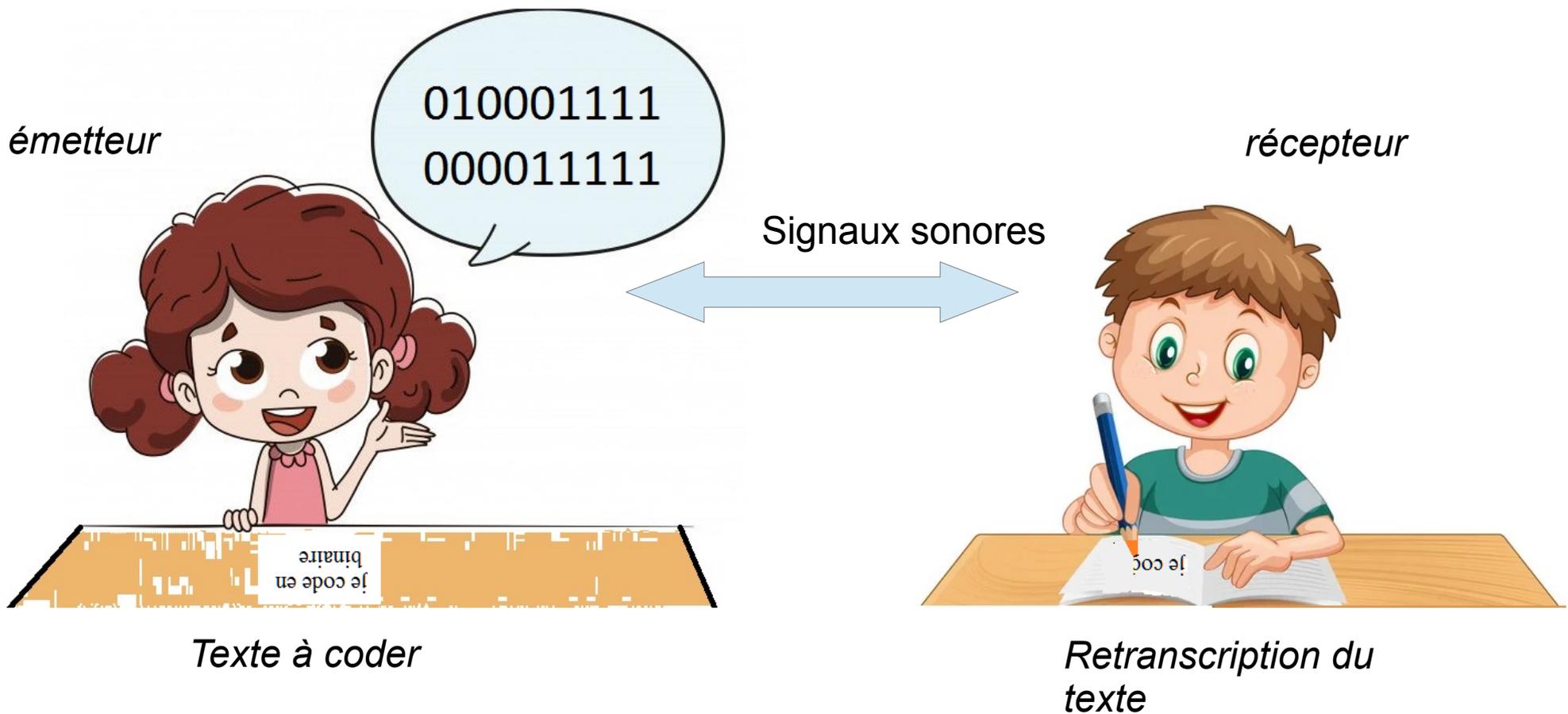
Question - problème

Comment coder du texte en binaire ?

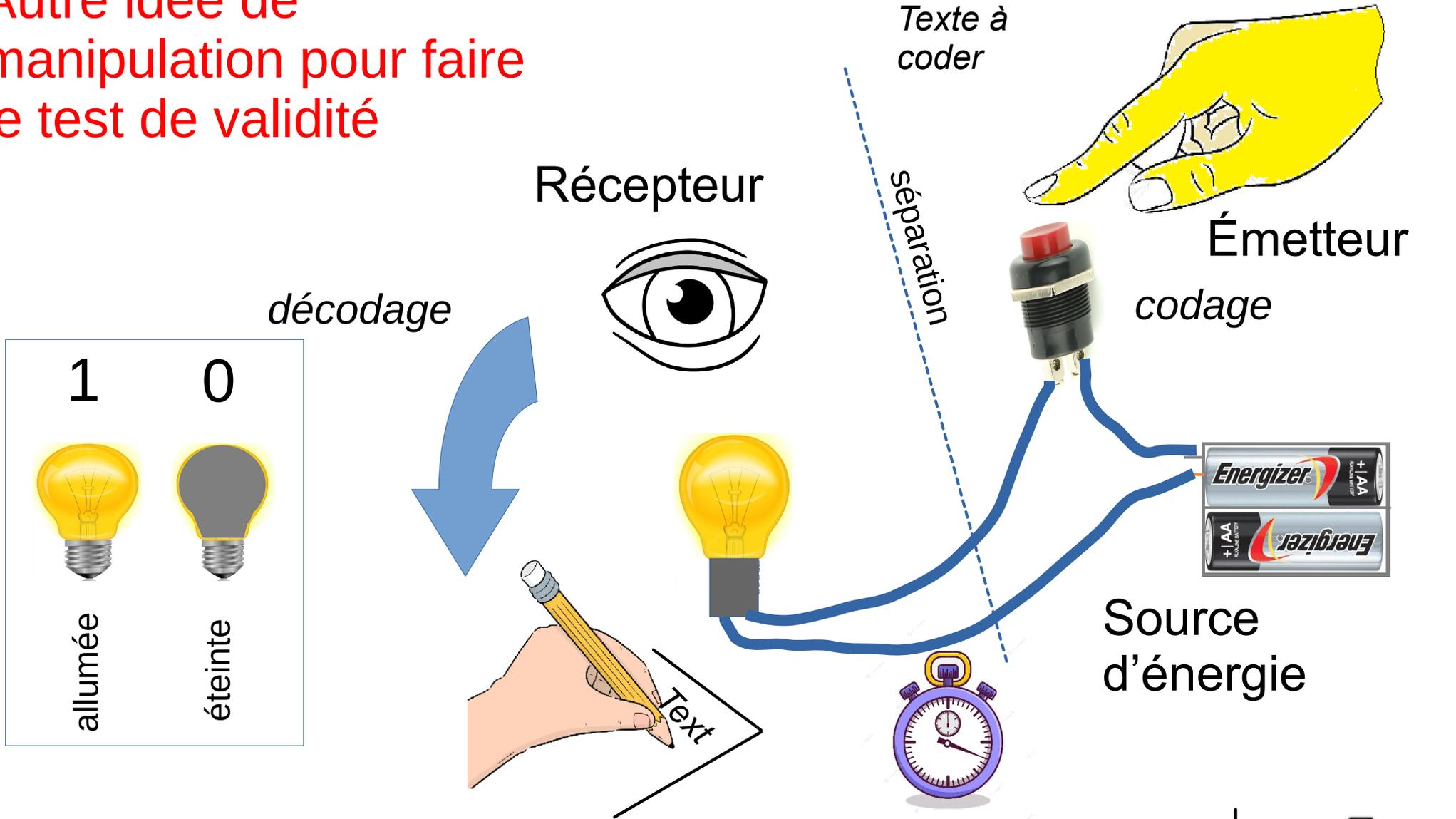
Hypothèses - Conjoncture

Indices : Activité 1 de manipulations de fichiers textes simples
 Qu'indique le suffixe « octo » (octogone) ?

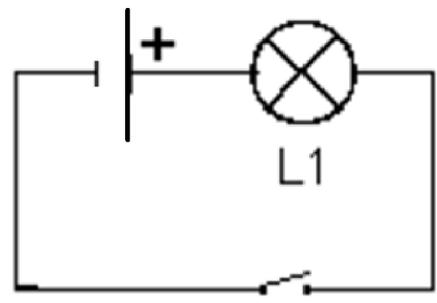
expérience *Retirer le texte à coder auprès du prof* (non connu avant l'expérience)



Autre idée de manipulation pour faire le test de validité



on peut exprimer le montage par un schéma



Quelques idées

Réponse de l'informatique

Chaque caractère typographique a son codage

e → 01100101

Le même nombre de bit pour tous : 8 à la suite



1 caractère : 8 Bit

↓
Permet de coder 256 signes différents

C'est pour cela qu'on a défini l'octet = 8 bit

1 octet correspond à l'enregistrement d'un caractère typographique

Permet de coder tous les symboles textuels

Alphabet latin :

a,b,c,d,.....z
+ A,B,C,D.....Z
+ é,à,è: ù,
+ 0,1,2,3...9
+ <, >, +, =
+ « ; ! §

→ Plein de symboles différents

Des tables associent le codage binaire aux symboles pour que les ordinateurs se comprennent.

Problème il y en a beaucoup, ils évoluent....

Une des plus utilisées : ISO/CEI 8859-15
En 1998 (symbole €)

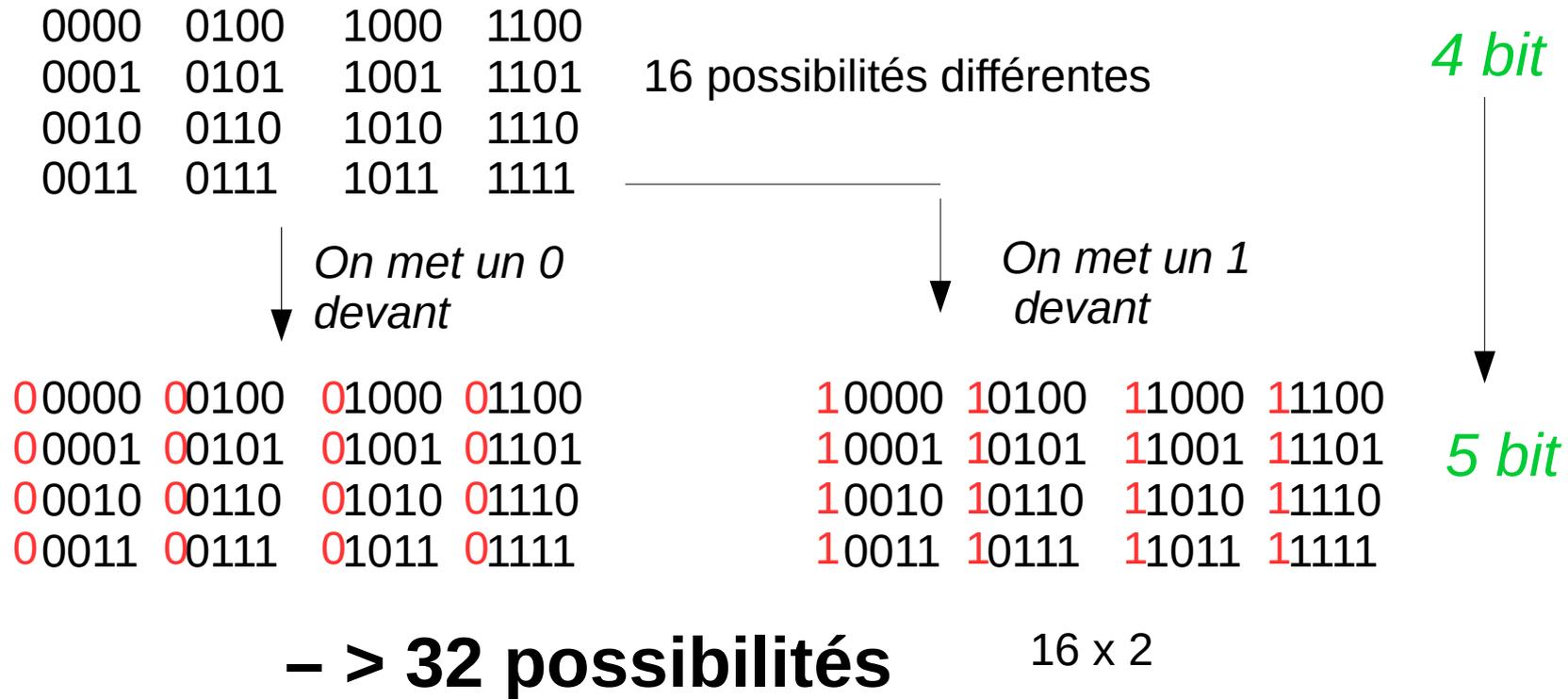
Il n'y a pas d'accent en langue anglaise, au début le codage des lettres était sur 7 bits ...

Comprendre l'intérêt de regrouper les bits

| | | | | | | | |
|---------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|-----------------------------|--------------------------------------|--------|
| Une information codée par 1 bit | 0 1 | 2 mots différents | Noir blanc | a b | | | |
| Une information codée par 2 bit | 00 01 | 10 11 | 4 possibilités | Noir rouge | bleu blanc | a b | c d |
| Une information codée par 3 bit | 000 001 | 010 011 | 8 possibilités | Noir rouge | bleu blanc | a b | c d |
| | 100 101 | 110 111 | | | | | |
| Une information codée par 4 bit | 0000 0001 0010 0011 | 0100 0101 0110 0111 | 1000 1001 1010 1011 | 1100 1101 1110 1111 | 16 possibilités différentes | <i>16 significations différentes</i> | |

Combien de possibilités différentes pouvons nous coder avec une information sur 5 bit ?

Combien de possibilités différentes pouvons nous coder avec une information sur 5 bit ?



Combien de possibilités différentes pouvons nous coder avec une information sur 6 bit ?

Combien de possibilités différentes pouvons nous coder avec une information sur 6 bit ?

5 bit – > **32 possibilités**

6 bit **32 x 2** – > **64 possibilités**

Combien de possibilités différentes pouvons nous coder avec une information sur 8 bit ?

7 bit **64 x 2** – > **128 possibilités**

8 bit **128 x 2** – > **256 possibilités**

Chaque « groupe » de 8 bits indique 1 des 250 caractères

11101001 ⇔ é 00100000 ⇔ espace

0101100 ⇔ ,

Plus petite unité d'information

binary Unit

Chiffre binaire

« Une lettre de l'ordinateur » : **1 Bit**

Une information numérique codée sur 16 Bit

0110101011000101

En Fait :

Unité informatique « usuelle »

Octet = 8 bits

Byte en Anglais

Une information numérique codée sur 2 octets

0110101011000101

Table ASCII sur 7 bits (128 signes)

| Bin | Char | Bin | Char | Bin | Char | Bin | Char |
|-----------|-------|-----------|-------|-----------|------|-----------|-------|
| 0000 0000 | [NUL] | 0010 0000 | space | 0100 0000 | @ | 0110 0000 | ` |
| 0000 0001 | [SOH] | 0010 0001 | ! | 0100 0001 | A | 0110 0001 | a |
| 0000 0010 | [STX] | 0010 0010 | " | 0100 0010 | B | 0110 0010 | b |
| 0000 0011 | [ETX] | 0010 0011 | # | 0100 0011 | C | 0110 0011 | c |
| 0000 0100 | [EOT] | 0010 0100 | \$ | 0100 0100 | D | 0110 0100 | d |
| 0000 0101 | [ENQ] | 0010 0101 | % | 0100 0101 | E | 0110 0101 | e |
| 0000 0110 | [ACK] | 0010 0110 | & | 0100 0110 | F | 0110 0110 | f |
| 0000 0111 | [BEL] | 0010 0111 | ' | 0100 0111 | G | 0110 0111 | g |
| 0000 1000 | [BS] | 0010 1000 | (| 0100 1000 | H | 0110 1000 | h |
| 0000 1001 | [TAB] | 0010 1001 |) | 0100 1001 | I | 0110 1001 | i |
| 0000 1010 | [LF] | 0010 1010 | * | 0100 1010 | J | 0110 1010 | j |
| 0000 1011 | [VT] | 0010 1011 | + | 0100 1011 | K | 0110 1011 | k |
| 0000 1100 | [FF] | 0010 1100 | , | 0100 1100 | L | 0110 1100 | l |
| 0000 1101 | [CR] | 0010 1101 | - | 0100 1101 | M | 0110 1101 | m |
| 0000 1110 | [SO] | 0010 1110 | . | 0100 1110 | N | 0110 1110 | n |
| 0000 1111 | [SI] | 0010 1111 | / | 0100 1111 | O | 0110 1111 | o |
| 0001 0000 | [DLE] | 0011 0000 | 0 | 0101 0000 | P | 0111 0000 | p |
| 0001 0001 | [DC1] | 0011 0001 | 1 | 0101 0001 | Q | 0111 0001 | q |
| 0001 0010 | [DC2] | 0011 0010 | 2 | 0101 0010 | R | 0111 0010 | r |
| 0001 0011 | [DC3] | 0011 0011 | 3 | 0101 0011 | S | 0111 0011 | s |
| 0001 0100 | [DC4] | 0011 0100 | 4 | 0101 0100 | T | 0111 0100 | t |
| 0001 0101 | [NAK] | 0011 0101 | 5 | 0101 0101 | U | 0111 0101 | u |
| 0001 0110 | [SYN] | 0011 0110 | 6 | 0101 0110 | V | 0111 0110 | v |
| 0001 0111 | [ETB] | 0011 0111 | 7 | 0101 0111 | W | 0111 0111 | w |
| 0001 1000 | [CAN] | 0011 1000 | 8 | 0101 1000 | X | 0111 1000 | x |
| 0001 1001 | [EM] | 0011 1001 | 9 | 0101 1001 | Y | 0111 1001 | y |
| 0001 1010 | [SUB] | 0011 1010 | : | 0101 1010 | Z | 0111 1010 | z |
| 0001 1011 | [ESC] | 0011 1011 | ; | 0101 1011 | [| 0111 1011 | { |
| 0001 1100 | [FS] | 0011 1100 | < | 0101 1100 | \ | 0111 1100 | |
| 0001 1101 | [GS] | 0011 1101 | = | 0101 1101 |] | 0111 1101 | } |
| 0001 1110 | [RS] | 0011 1110 | > | 0101 1110 | ^ | 0111 1110 | ~ |
| 0001 1111 | [US] | 0011 1111 | ? | 0101 1111 | | 0111 1111 | [DEL] |

extensions sur 8 bit (Octet)

Des signes en plus... signes)

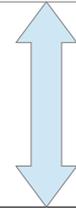
| | | | | | | | | | | | | | |
|----|----|----|----|----|-----|-----|-----|-----|---|-----|---|-----|---|
| 0 | 32 | 64 | @ | 96 | ' | 128 | 160 | 192 | À | 224 | à | | |
| 1 | 33 | ! | 65 | A | 97 | a | 129 | 161 | ¡ | 193 | Á | 225 | á |
| 2 | 34 | " | 66 | B | 98 | b | 130 | 162 | ¢ | 194 | Â | 226 | â |
| 3 | 35 | # | 67 | C | 99 | c | 131 | 163 | £ | 195 | Ã | 227 | ã |
| 4 | 36 | \$ | 68 | D | 100 | d | 132 | 164 | ¤ | 196 | Ä | 228 | ä |
| 5 | 37 | % | 69 | E | 101 | e | 133 | 165 | ¥ | 197 | Å | 229 | å |
| 6 | 38 | & | 70 | F | 102 | f | 134 | 166 | ¦ | 198 | Æ | 230 | æ |
| 7 | 39 | ' | 71 | G | 103 | g | 135 | 167 | § | 199 | Ç | 231 | ç |
| 8 | 40 | (| 72 | H | 104 | h | 136 | 168 | ¨ | 200 | È | 232 | è |
| 9 | 41 |) | 73 | I | 105 | i | 137 | 169 | © | 201 | É | 233 | é |
| 10 | 42 | * | 74 | J | 106 | j | 138 | 170 | ª | 202 | Ê | 234 | ê |
| 11 | 43 | + | 75 | K | 107 | k | 139 | 171 | « | 203 | Ë | 235 | ë |
| 12 | 44 | , | 76 | L | 108 | l | 140 | 172 | ¬ | 204 | Ì | 236 | ì |
| 13 | 45 | - | 77 | M | 109 | m | 141 | 173 | - | 205 | Í | 237 | í |
| 14 | 46 | . | 78 | N | 110 | n | 142 | 174 | ® | 206 | Î | 238 | î |
| 15 | 47 | / | 79 | O | 111 | o | 143 | 175 | - | 207 | Ï | 239 | ï |
| 16 | 48 | 0 | 80 | P | 112 | p | 144 | 176 | ° | 208 | Ð | 240 | ð |
| 17 | 49 | 1 | 81 | Q | 113 | q | 145 | 177 | ± | 209 | Ñ | 241 | ñ |
| 18 | 50 | 2 | 82 | R | 114 | r | 146 | 178 | ² | 210 | Ò | 242 | ò |
| 19 | 51 | 3 | 83 | S | 115 | s | 147 | 179 | ³ | 211 | Ó | 243 | ó |
| 20 | 52 | 4 | 84 | T | 116 | t | 148 | 180 | ´ | 212 | Ô | 244 | ô |
| 21 | 53 | 5 | 85 | U | 117 | u | 149 | 181 | µ | 213 | Õ | 245 | õ |
| 22 | 54 | 6 | 86 | V | 118 | v | 150 | 182 | ¶ | 214 | Ö | 246 | ö |
| 23 | 55 | 7 | 87 | W | 119 | w | 151 | 183 | · | 215 | × | 247 | ÷ |
| 24 | 56 | 8 | 88 | X | 120 | x | 152 | 184 | , | 216 | Ø | 248 | ø |
| 25 | 57 | 9 | 89 | Y | 121 | y | 153 | 185 | ¸ | 217 | Ù | 249 | ù |
| 26 | 58 | : | 90 | Z | 122 | z | 154 | 186 | ¸ | 218 | Ú | 250 | ú |
| 27 | 59 | ; | 91 | [| 123 | { | 155 | 187 | » | 219 | Û | 251 | û |
| 28 | 60 | < | 92 | \ | 124 | | 156 | 188 | ¼ | 220 | Ü | 252 | ü |
| 29 | 61 | = | 93 |] | 125 | } | 157 | 189 | ½ | 221 | Ý | 253 | ý |
| 30 | 62 | > | 94 | ^ | 126 | ~ | 158 | 190 | ¾ | 222 | Þ | 254 | þ |
| 31 | 63 | ? | 95 | _ | 127 | | 159 | 191 | ¿ | 223 | ß | 255 | ÿ |

Exemple : codage d'un texte en binaire

```
0101011001101111011101010111001100100000011000110110
1000011000010110111001110100011010010110010101111010
0011111100100000011010100101110000100111011001010110
1110001000000111001101110101011010010111001100100000
0110011001101111011100100111010000100000011000010110
100101110011011001010000110100001010010001010111010
000100000011000100110100101100101011011100010000100
1000000110010001100001011011100111001101100101011110
1000100000011011010110000101101001011011100111010001
1001010110111001100001011011100111010000101110
```

512 bits

64 octets



Codage ISO/CEI 8859-15

11101001 ↔ é

00111111 ↔ ?

00100000 ↔ espace

Codage Ascii étendu

Vous chantiez? j'en suis fort aise :
Et bien! dansez maintenant.

64 signes...

Texte à taper avec

Le « bloc note » :



Enregistrement d'un texte

Texte

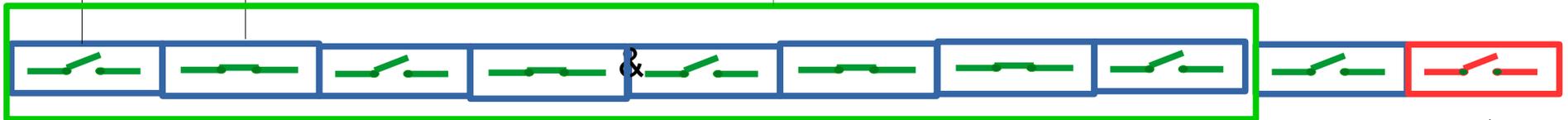
Vous chantez?....

Traduction en binaire

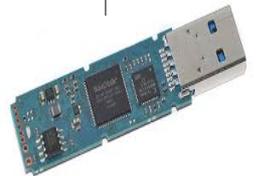
01010110 011011110111010101110011001000
00011000110110100001100001011011100111
01000110100101100101011110100011111100
10000

Enregistrement du V

0 1 0 1 0 1 1 0



1 lettre V = 8 BIT = 1 Octet



Capacité clef USB

4 Go = 4000000000 octets = 32000000000 bit

On pourrait enregistrer

4 milliards de lettres ---> ~ 3000 livres...

Questions bilan A)

Feuille du classeur (culture informatique)

Cycle 4

Culture informatique

nom prénom

Qu'est ce qu'un fichier informatique ?

A) principe du codage binaire

Le langage de l'ordinateur ne comprend que deux symboles que l'on note habituellement ___ et ___.

Toutes les informations sont donc codées à l'aide de ces deux symboles. On dit que le langage est _____

Chaque «symbole» est appelé _____

```
010100011010000110100101100110010
01101011110110110101001001000000
1011101101101001000000111010001100
1101011101101000101101100110011001
0000110010101110010011001010010000
00110011011001010111010001110100011
100011010000110010100100000011000
01111001001000001101100110000101
0000110111101100110001000000110100
00110100101110011001000000110100101
10101001010010000001100100110000
01000001110100011001011110000111
0101101110011001110010000011011101
1100100110010100100000011001100110
01011010001101010001100101011100
```

nom prénom

Qu'est ce qu'un fichier informatique?

A) principe du codage binaire

Le langage de l'ordinateur ne comprend que deux symboles que l'on note habituellement 0 et 1.

Toutes les informations sont donc codées à l'aide de ces deux symboles. On dit que le langage est binaire

Chaque «symbole» est appelé bit

```
0101000110100001101001011100110010
011011011110110110101100101001000001
10111101101101001000000111010001100
110110111011010010110111001100111001
0000110010101110010011001010010001
001100111011001010111010001110100011
100011010000110010100100000011000
0111100100100000011101100110000101
00000110111101100110001000000110100
0011010010111001100100000110100101
1010110010100100000011100100110001
01000000111010001100101011110000111
0101101110011001110010000001101110
1100100110010100100000011001100110
01011101000111010001101001011011100
```

Question bilan B

B) codage du texte

Dans l'application « bloc note » on peut taper du texte sans mise en forme

Voici le texte tapé : *Voici un exemple du codage d'un texte en binaire*

Voici son codage binaire , suivant les norme ASCII (Un caractère est codé avec 8 bits)

```
01010110 01101111 01101001 01100011 01101001 00100000 01110101
01101110 00100000 01100101 01111000 01100101 01101101 01110000
01101100 01100101 00100000 01100100 01110101 00100000 01100011
01101111 01100100 01100001 01100111 01100101 00100000 01100100
00100111 01110101 01101110 00100000 01110100 01100101 01111000
01110100 01100101 00100000 01100101 01101110 00100000 01100010
01101001 01101110 01100001 01101001 01110010 01100101
```

1) Quel est le codage correspondant à la lettre « V » (majuscule) ?

2) Quel est le codage correspondant à la lettre « e » ? _____

3) Quelle taille devrait avoir ce fichier texte enregistré avec l'application « bloc note » (sans en tete)

a) En bits ? _____

b) En octets ? _____

Question bila n 2

B) codage du texte

Dans l'application « bloc note » on peut taper du texte sans mise en forme

Voici le texte tapé : **Voici un exemple du codage d'un texte en binaire**

Voici son codage binaire , suivant les norme ASCII (Un caractère est codé avec 8 bits)

01010110 01101111 01101001 01100011 01101001 **00100000** 01110101
01101110 **00100000** 01100101 01111000 01100101 01101101 01110000
01101100 01100101 00100000 01100100 01110101 00100000 01100011
01101111 01100100 01100001 01100111 01100101 00100000 01100100
00100111 01110101 01101110 00100000 01110100 01100101 01111000
01110100 01100101 00100000 01100101 01101110 00100000 01100010
01101001 01101110 01100001 01101001 01110010 **01100101**

1) Quel est le codage correspondant à la lettre « V » (majuscule) ?

01010110

2) Quel est le codage correspondant à la lettre « e » ? 01100101

3) Quel est le codage correspondant à un espace : **00100000**

4) Quelle taille devrait avoir ce fichier texte enregistré avec l'application « bloc note » (sans en tete)

a) En bits ? 384 bits

b) En octets ? **48 (il y a 48 caractères)**

Fichier enregistré dans le disque dur

Les données (codage binaire du texte) sont chargées dans la mémoire vive

Le processeur traite les données et envoie les signaux pour allumer les pixels de l'écran...



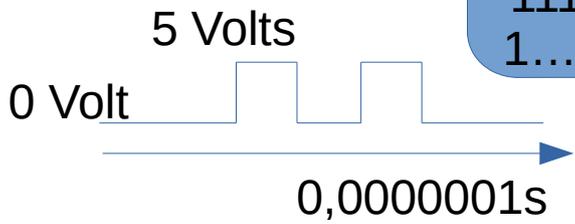
Cf : Petite anatomie d'unités de stockage



0011111111
1111111111
1111111111
111100111
1.....

0011111111
1111111111
1111111111
111100111
1.....

0011111111
1111111111
1111111111
111100111
1.....



Des signaux électriques dans le temps....

Le texte s'affiche sur l'écran

Cf : les périphériques traitant les images

