



Travaux dirigés n° 2

Représentation des données

Au cœur de l'ordinateur se trouve une mémoire contenant des (représentations binaires d') entiers et une unité de calcul capable d'effectuer les opérations arithmétiques sur ces entiers. Or chacun sait que les ordinateurs peuvent aussi travailler sur des textes, des images, du son, ... et même se communiquer ces données à travers un réseau. Cela suppose (au moins) deux choses : d'une part, on peut coder le texte, l'image, le son ... par une suite d'entiers (problème mathématique et physique non évident pour le son, par exemple!) ; d'autre part, il faut que tout le monde se mette d'accord sur ces codages, en adoptant des normes.

Vous trouverez en annexe des informations sur certaines organisations internationales de normalisation.

On prendra comme unité de base l'octet : suite de 8 bits représentant un entier compris entre 0 et 255. On notera le contenu d'un octet par sa valeur exprimée en hexadécimal (2 chiffres hexadécimaux).

Le texte

Un texte étant une suite de caractères, on peut représenter un texte par la suite des représentations des caractères qui le constituent. Le principe est alors le suivant : on se met d'accord sur l'ensemble des caractères à représenter, on assigne un numéro (entier) à chacun de ces caractères puis on définit une méthode pour associer à ce numéro un octet.

ISO 646

Cette norme reprend le code **ASCII** *American Standard Code for Information Interchange* qui définit un ensemble de 128 caractères numérotés de 0 à 127. On code le caractère par l'octet correspondant à son numéro (voir en annexe la table de codage ASCII en ISO 646 : a).

NOTATIONS : $chr(n)$ pour désigner le caractère de numéro n et $ord(c)$ pour le numéro du caractère c .

Exercice 1 (Codage et décodage)

1°) Déterminez les caractères de codes :

41	57	34	5a	3d	5f

2°) Décodez le texte suivant :

4a 27 61 69 20 70 65 72 64 75 20 6c 61 20 74 65 74 65 2c 20 64 65 70 75 69 73 20 71 75 65 20 6a 27
61 69 20 76 75 20 53 75 7a 65 74 74 65

3°) Codez le texte suivant : Il est impossible de se tromper !

4°) Pourquoi les français ne sont-ils pas satisfaits de cette norme?

Exercice 2 (Classement alphabétique)

Classez par ordre alphabétique croissant les trois textes : "Dupont", "dupont", "Du Pont"

Exercice 3 (Un peu d'algorithmique)

Décrivez comment on peut réaliser chacune des opérations suivantes :

1. Le texte est en minuscules et on désire le transformer en majuscules.
2. Le texte est un message secret et on désire le crypter en décalant chaque lettre de -4 positions.
Par exemple : "ausecoursvercingetorixattaque" doit devenir : "wqoaykqnoranyejcapknetwppwmqa"
3. On souhaite supprimer tous les espaces.
4. On souhaite doubler tous les espaces.

N.B. : en algorithmique, une chaîne de caractères est vue comme un tableau. Aussi, $longueur(t)$ donne la longueur de la chaîne et $t[i]$ permet d'accéder au i ème caractère de t (avec i compris entre 0 et $longueur(t)-1$).

ISO 8859-1

Avec un octet on peut coder 256 caractères, on a donc défini plusieurs extensions du codage **ASCII** en ajoutant 128 caractères pour mieux couvrir les besoins des différentes langues. Par exemple :

ISO 8859-1 pour les langues de l'ouest européen,

ISO 8859-2 pour le polonais, le tchèque, ...

etc ...

Voir en annexe la table de codage ASCII en ISO 8859-1 (b) pour le jeu le plus utilisé.

Exercice 4 (Les accents)

1°) Vérifiez que la transformation minuscule vers majuscule est correctement étendue aux caractères accentués.

2°) Pourquoi les français ne sont-ils toujours pas satisfaits par ce codage ?

Les valeurs numériques : entiers**Entiers non signés sur 2, 4, 8 octets**

$2902 = 11 * 256 + 86 = (b56)_{16}$ peut se coder, sur deux octets : $[0b\ 56]_{16}$ ou $[0000\ 1011\ 0101\ 0110]_2$

Exercice 5 (Représentation d'entiers)

1°) Donnez le codage sur 4 octets de la valeur 5165.

2°) Codez sur deux octets les valeurs : 57, 786, 4503, 57112, ...

3°) Quelle est la valeur représentée sur un octet par $[c9]_{16}$? par $[1101\ 1011]_2$?

Exercice 6 (Représentation)

On considère la suite d'octets suivante : 5665726c61696e65

1°) Décodez-la en supposant qu'il s'agit d'un texte codé en **ASCII**.

2°) Idem, mais en supposant qu'il s'agit d'une suite de 8 entiers non signés codés sur un octet.

Entiers signés

On considère ici un seul octet. Il existe plusieurs solutions :

♣ **signe et valeur absolue** le bit de poids fort code le signe (0 pour + et 1 pour -); les autres bits codent la valeur absolue.

- $(17)_{10} \leftrightarrow [11]_{sva16}$
- $-(17)_{10} \leftrightarrow [11]_{sva16} + [80]_{sva16} = [91]_{sva16}$

◇ **biais** le codage s'obtient en ajoutant une constante (généralement $(128)_{10} = (80)_{16}$) *aux positifs et aux négatifs*

- $(17)_{10} \leftrightarrow [91]_{b16}$ car $(80)_{16} + (11)_{16} = (91)_{16}$: $128 + 17 = 145 = (91)_{16}$
- $-(17)_{10} \leftrightarrow [6f]_{b16}$ car $(80)_{16} - (11)_{16} = (6f)_{16}$: $128 - 17 = 111 = (6f)_{16}$

♡ ...

♠ **complément à 2** le codage s'obtient en ajoutant $(256)_{10} = (100)_{16}$ *uniquement aux négatifs*

- $(17)_{10} \leftrightarrow [11]_{c16}$
- $-(17)_{10} \leftrightarrow [ef]_{c16}$ car $(100)_{16} - (11)_{16} = (ef)_{16}$: $256 - 17 = 239 = (ef)_{16}$

Exercice 7 (Les opposés)

1°) Expliquez pourquoi dans les deux derniers systèmes, on a toujours : $code(n) + code(-n) = (256)_{10} = (100)_{16}$

2°) On définit l'opération notée \sim par le tableau suivant (à compléter) :

x	0	1	2	3	...	e	f
$\sim x$	f	e					0
$x + \sim x =$	f	f	f	f	...	f	f

Vérifiez que $code(-n) = \sim code(n) + 1$

3°) Calculez la représentation en complément à 2 sur un octet des entiers : -1, -45, 76, -87, ...

4°) A quels entiers correspondent les codages suivants (en complément à 2 sur 8 bits) : $[1011\ 0010]_{c2}$, $[0100\ 1110]_{c2}$, $[1111\ 1101]_{c2}$, $[7c]_{c16}$ et $[f5]_{c16}$?

Exercice 8 (Les différents systèmes)

1°) Complétez le tableau suivant :

octet	Valeur codée en signe et VA	Valeur codée avec biais de 128	Valeur codée en complément à 2
00			
:			
7f			
80			
81			
:			
ff			

2°) Justifiez l'appellation "bit de signe" donnée au bit de poids fort.

3°) Comparez les opérations : $3 + (-4)$ et $3 - (-4)$ dans chacun des 3 systèmes.

4°) Comparez l'opération : $3 * (-4)$ dans les deux derniers systèmes.

Exercice 9 (Le complément)

1°) Codez en complément à 2 sur 2 octets : 45, -117, 678, -2764, ...

2°) On considère le nombre codé sur un octet en complément à 2 par $[65]_{c16}$: donnez son codage sur 2 octets, 4 octets. De même pour $[97]_{c16}$.

Opérations et (?) débordements**Exercice 10 (Quelques opérations)**

On considère des entiers signés codés sur 1 octet.

1°) Codez 95, 56 et -117

2°) Effectuez en binaire l'opération $95 + 56$; vérifiez le résultat.

3°) Même chose avec l'opération binaire correspondant à $56 - (-117)$.

4°) Même chose avec l'opération binaire correspondant à $95 * 56$.

5°) Vous avez sans doute utilisé l'écriture binaire de la représentation : et avec l'écriture octale ? et avec l'écriture hexadécimale ?

ANNEXES

Quelques organisations internationales de normalisation et la façon dont elles se présentent :

ISO *Si les normes n'existaient pas, nous le remarquerions sans tarder. La contribution des normes, bien qu'elle soit très souvent imperceptible, est considérable dans la plupart des aspects de nos vies. L'importance des normes se manifeste lorsque celles-ci font défaut. Ainsi, en tant qu'acheteur ou utilisateur de produits, nous remarquons rapidement si ces produits sont de mauvaise qualité, s'ils ne sont pas adaptables, ou s'ils sont incompatibles avec d'autres appareils que nous possédons déjà, s'ils ne sont pas fiables ou présentent des dangers. Il nous semble aller de soi que les produits répondent à nos attentes. [...]*

ISO (Organisation internationale de normalisation) est le plus grand organisme de normalisation au monde. L'ISO a pour activité principale l'élaboration de normes techniques, mais ces dernières ont aussi d'importants aspects économiques et sociaux. Les normes ISO ont une influence positive, non seulement pour les ingénieurs et les fabricants, auxquels elles apportent des solutions à des problèmes fondamentaux de production et de distribution, mais pour la société dans son ensemble...

ITU/UIT *The International Telecommunication Union (ITU) is the leading publisher of telecommunication technology, regulatory and standard information. Over 4'000 titles in printed form and growing number on CD-ROM and Online...*

IEC/CEI *The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes international standards for all electrical, electronic and related technologies...*

	000	001	002	003	004	005	006	007
0	[NUL]	[DLE]	[SP]	0	@	P	'	p
1	[SOH]	[DC1]	!	1	A	Q	a	q
2	[STX]	[DC2]	"	2	B	R	b	r
3	[ETX]	[DC3]	#	3	C	S	c	s
4	[EOT]	[DC4]	\$	4	D	T	d	t
5	[ENQ]	[NAK]	%	5	E	U	e	u
6	[ACK]	[SYN]	&	6	F	V	f	v
7	[BEL]	[ETB]	'	7	G	W	g	w
8	[BS]	[CAN]	(8	H	X	h	x
9	[HT]	[EM])	9	I	Y	i	y
A	[LF]	[SUB]	*	:	J	Z	j	z
B	[VT]	[ESC]	+	;	K	[k	{
C	[FF]	[FS]	,	<	L	\	l	
D	[CR]	[GS]	-	=	M]	m	}
E	[SO]	[RS]	.	>	N	^	n	~
F	[SI]	[US]	/	?	O	_	o	[DEL]

(a) ISO 646

	008	009	00a	00b	00c	00d	00e	00f
0	[XX̄]	[DC̄S]	[NB̄S]	°	À	Đ	à	đ
1	[XX̄]	[PŪ1]	ı	±	Á	Ñ	á	ñ
2	[SP̄H]	[PŪ2]	ç	²	Â	Ò	â	ò
3	[NB̄H]	[ST̄S]	£	³	Ã	Ó	ã	ó
4	[IND̄]	[CC̄H]	¤	´	Ä	Ô	ä	ô
5	[NEL̄]	[MW̄]	¥	µ	Å	Õ	å	õ
6	[SS̄A]	[SP̄A]	ı	¶	Æ	Ö	æ	ö
7	[ES̄A]	[EP̄A]	§	·	Ç	×	ç	÷
8	[HT̄S]	[SŌS]	¨	,	È	Ø	è	ø
9	[HT̄J]	[XX̄]	©	¹	É	Ù	é	ù
A	[VT̄S]	[SC̄I]	ª	º	Ê	Ú	ê	ú
B	[PL̄D]	[CS̄I]	«	»	Ë	Û	ë	û
C	[PL̄U]	[ST̄]	¬	¼	Ì	Ü	ì	ü
D	[RĪ]	[OS̄C]	[SH̄T]	½	Í	Ý	í	ý
E	[SS̄2]	[PM̄]	®	¾	Î	Þ	î	þ
F	[SS̄3]	[AP̄C]	™	ı	Ï	ß	ï	ÿ

(b) ISO 8859-1