



Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Codage des nombres et des caractères



Problématique

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Le système de représentation des nombres le plus naturel est le système décimal. (0,1,2,3,4,5,6,7,8,9)

Problème : l'ordinateur code toute information en langage binaire (succession de 1 et 0)

Comment coder 74, 3.23 ou "bonjour!" en binaire ?



Vocabulaire

Codage des
nombres et
des
caractères

bit : binary digit = Unité d'information pouvant valoir
1 ou 0

octet (o) : ensemble de 8 bits (en anglais Byte (B))

Type :	Disque local	
Système de fichiers :	NTFS	
<hr/>		
 Espace utilisé :	306 892 697 600 octets	285 Go
 Espace libre :	39 990 919 168 octets	37,2 Go
<hr/>		
Capacité :	346 883 616 768 octets	323 Go



Vocabulaire

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Préfixes système décimal

Nom	symbole	Valeur	Exemple
kilo	k	10^3	1 ko = 1×10^3 o
Méga	M	10^6	1 Mo = 1×10^6 o
Giga	G	10^9	1 Go = 1×10^9 o
Tera	T	10^{12}	1 To = 1×10^{12} o



Vocabulaire

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Préfixes système binaire

Nom	symbole	Valeur	Exemple
kibi	ki	2^{10}	1 kio = 1024 o
Mébi	Mi	2^{20}	1 Mio = 1024 kio = 1 048 576 o
Gibi	Gi	2^{30}	1 Gio = 1 099 511 627 776 o
Tebi	Ti	2^{40}	1 Tio = 1 125 899 906 842 624 o



Vocabulaire

Codage des
nombres et
des
caractères


Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Général Emblèmes Permissions Ouvrir avec Notes

 **Nom :** HTC Android Phone: Système de fichiers 7,9 GB

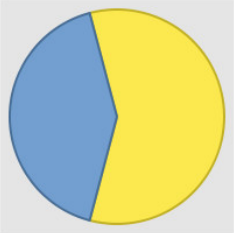
Type : dossier (inode/directory)

Contenu : 1 014 éléments

Emplacement : computer:///

Volume : Système de fichiers 7,9 GB

Espace libre : 3,1 Gio



■ 4,3 Gio utilisé

■ 3,1 Gio libre

Capacité totale : 7,4 Gio

Type de système de fichiers : msdos

Aide Fermer



Base N vers base décimale

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

- ▶ Décimale : $B_{10} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

Exemple : $(1024)_{10} = 1 \times 10^3 + 0 \times 10^2 + 2 \times 10^1 + 4 \times 10^0$

- ▶ Binaire : $B_2 = \{0, 1\}$

Exemple : $(1101)_2 = 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$

- ▶ hexadécimal :

$$B_{16} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F\}$$

Exemple :

$$(4E13)_{16} = 4 \times 16^3 + 14 \times 16^2 + 1 \times 16^1 + 3 \times 16^0$$

- ▶ Base N : $B_N = \{0, 1, \dots, N\}$

Exemple : $(wxyz)_N = w \times N^3 + x \times N^2 + y \times N^1 + z \times N^0$

On peut facilement passer d'un nom entier positif initialement écrit en base N à son écriture en base 10



Les 3 principales bases de numération

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

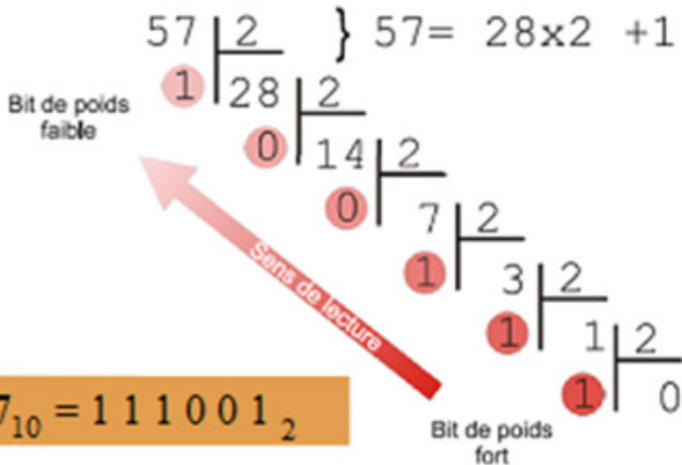
Codage des
caractères

Base 10	Base 2	Base 16
0	0	0
1	1	1
2	10	2
3	11	3
4	100	4
5	101	5
6	110	6
7	111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F



Passage de la base 10 à la base 2

On procède par divisions euclidiennes successives.





Passage de la base 10 à la base 16

Codage des
nombres et
des
caractères

Bases de
numération

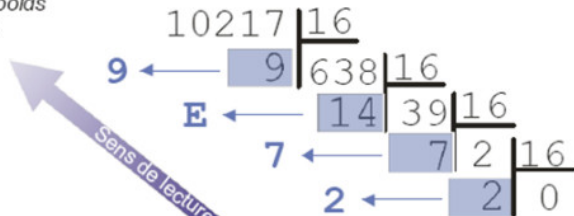
Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

On procède par divisions euclidiennes successives.

*Chiffre de poids
faible*



Sens de lecture

*Chiffre de poids
fort*

$$10217_{10} = 27E9_{16}$$



Codage des nombres entiers

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Comment stocker les nombres entiers (positifs ou négatifs) en mémoire ?

- ▶ Première idée : Réserver le bit de poids fort pour le signe.

Ainsi, sur un octet (8 bits) $(13)_{10}$ s'écrirait $(00001101)_2$ et $(-13)_{10}$ s'écrirait alors $(10001101)_2$.

Inconvénients :

- ▶ zéro n'a pas une écriture unique : $(00000000)_2$ et $(10000000)_2$;
- ▶ l'addition classique ne fonctionne pas pour les nombres négatifs. ($6-3=?$)



Complément à deux

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Pour palier les problèmes cités précédemment les entiers naturels sont généralement codés par la méthode du complément à deux :

- ▶ les nombres positifs sont simplement représentés par leur écriture en base 2 ;
- ▶ Pour les nombres négatifs, on remplace chaque bit par son « complément à deux » : complément logique de chaque bit puis on ajoute 1. Pour un nombre a exprimé en décimal, le complément à deux de a codé sur N bits binaires vaut $(2^N - |a|)_2$
- ▶ Le bit de poids le plus fort est toujours réservé pour coder le signe.

Ainsi, sur un octet, $(13)_{10}$ s'écrit $(00001101)_2$ et $(-13)_{10}$ s'écrit $(11110011)_2$.



Complément à deux

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Avec cette méthode :

- ▶ on peut réaliser des opérations entre nombres entiers ;
- ▶ zéro a une unique écriture binaire ;
- ▶ on peut coder sur un octet des nombres a compris entre $(10000000)_2$ soit $a_{min} = (-128)_{10}$ et $(01111111)_2$ soit $a_{max} = (127)_{10}$ ($a \in [-2^N; 2^N - 1]$) avec N le nombre de bits)



Complément à deux

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Exemples :

▶ $-63+28=?$

▶ $63-28=?$

▶ $63-63=?$

▶ $103+65=?$

▶ $-103-65=?$



Des phénomènes étranges...

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

STRANGER THINGS



Vrai ou Faux ?

- ▶ $1/3 = 1 - 2/3$
- ▶ $2 - (2 + 10^{-20}) = (2 - 2) + 10^{-20}$
- ▶ $a = 9.0$ donc $(a^a)^a \approx 1.966 \times 10^{77}$
- ▶ $2 \times 10^{-324} > 1 \times 10^{-324}$
- ▶ $0,1 + 0,2 = 0,3$



Décomposition d'un nombre réel non entier sur une base binaire

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

▶ $(3,25)_{10} = 1 \times 2^1 + 1 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2}$ donc

$$(3,25)_{10} = (11,01)_2$$

▶ $(3,7)_{10} = (?)_2$

Conclusion :



Nombres à virgule flottante (float)

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

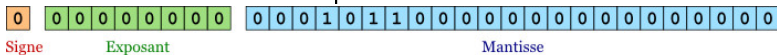
Codage des
nombres
réels

Codage des
caractères

La norme IEEE 754 définit comment l'ordinateur va coder de manière approchée les nombres réels (non entiers (int)). Il en existe deux versions :

- ▶ Le codage simple précision (sur 32 bits)
- ▶ Le codage double précision (sur 64 bits)

Un nombre flottant se compose de 3 informations :



$$a = (-1)^s m^e$$



Nombres à virgule flottante (float)

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

- ▶ Le bit de signe s :
0 : positif, 1 : négatif.
- ▶ La mantisse m :
 - ▶ nombre à virgule tel que $m \in [1, 2[$ codé sur 23 bits en simple précision (et 52 bits en double précision),
 - ▶ Comme la partie entière commence toujours par 1, celle-ci n'est pas codée. On ne code que les nombres après la virgule.
- ▶ L'exposant e^b (biaisé)
 - ▶ nombre entier codé sur 8 bits en simple précision (ou 11 bits en double précision),
 - ▶ L'exposant non biaisé (e) peut être positif ou négatif. Pour coder ce nombre, la règle est d'ajouter à l'exposant un biais (valeur constante) afin d'obtenir une valeur strictement positive : $e^b = e + 2^{N-1} - 1$ où N est le nombre de bits qui codent l'exposant (8 ou 11).
 - ▶ e_b est, bien entendu codé en binaire



Exemples : Comment coder $(3, 25)_{10}$ et $(3, 7)_{10}$ en flottant simple précision ?

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères



Passage d'un flottant à une approximation décimale

Codage des
nombres et
des
caractères

Déterminer l'écriture scientifique en base 10 du nombre flottant :
111000000101111000000000...
...00

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

[http ://www.binaryconvert.com/](http://www.binaryconvert.com/)



Quelques valeurs spéciales (en simple précision)

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

• Le plus grand nombre +	0 11111110 11111111111111111111111111111111
• Le plus grand nombre -	1 11111110 11111111111111111111111111111111
• Zéro +	0 00000000 000000000000000000000000000000
• Zéro -	1 00000000 000000000000000000000000000000
• Le plus petit nombre +	1 00000000 000000000000000000000000000001
• Le plus petit nombre -	1 00000001 000000000000000000000000000000
• Pi	0 10000000 10010010000111111011011
• + Infini	0 11111111 000000000000000000000000000000
• - Infini	1 11111111 000000000000000000000000000000
• Pas de valeur NaN (<i>Not A Number</i>)	1 11111111 100000000000000000000000000000



Tour d'horizon des problèmes rencontrés

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

- ▶ Dépassement de capacité (Overflow)
- ▶ Comparaison à zéro (Underflow)
- ▶ Absorption
- ▶ Cancellation



Codage des caractères

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Le codage des caractères est une convention qui permet, à travers un codage connu de tous, de transmettre de l'information textuelle, là où aucun support ne permet l'écriture scripturale.

Cela consiste à représenter chaque caractère, par un motif visuel, un motif sonore, ou une séquence abstraite. Les techniques des ordinateurs reposent sur l'association d'un caractère à un nombre.



Code ASCII

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Le code ASCII est l'un des plus anciens codes utilisés en informatique.

Le code ASCII (American Standard Code for Information Interchange) définit un jeu de 128 caractères numérotés de 0 à 127. Sept bits suffisent donc à les coder. Cependant, comme la plupart des ordinateurs travaillent sur des données représentées sur un nombre de bits multiple de huit, les caractères ASCII sont codés sur un octet, le bit de poids fort étant 0.



Code ASCII

Codage des nombres et des caractères

Bases de numération

Codage des nombres entiers

Codage des nombres réels

Codage des caractères

ASCII control characters

00	NULL	(Null character)
01	SOH	(Start of Header)
02	STX	(Start of Text)
03	ETX	(End of Text)
04	EOT	(End of Trans.)
05	ENQ	(Enquiry)
06	ACK	(Acknowledgement)
07	BEL	(Bell)
08	BS	(Backspace)
09	HT	(Horizontal Tab)
10	LF	(Line feed)
11	VT	(Vertical Tab)
12	FF	(Form feed)
13	CR	(Carriage return)
14	SO	(Shift Out)
15	SI	(Shift In)
16	DLE	(Data link escape)
17	DC1	(Device control 1)
18	DC2	(Device control 2)
19	DC3	(Device control 3)
20	DC4	(Device control 4)
21	NAK	(Negative acknowl.)
22	SYN	(Synchronous idle)
23	ETB	(End of trans. block)
24	CAN	(Cancel)
25	EM	(End of medium)
26	SUB	(Substitute)
27	ESC	(Escape)
28	FS	(File separator)
29	GS	(Group separator)
30	RS	(Record separator)
31	US	(Unit separator)
127	DEL	(Delete)

ASCII printable characters

32	space	64	@	96	`
33	!	65	A	97	a
34	"	66	B	98	b
35	#	67	C	99	c
36	\$	68	D	100	d
37	%	69	E	101	e
38	&	70	F	102	f
39	'	71	G	103	g
40	(72	H	104	h
41)	73	I	105	i
42	*	74	J	106	j
43	+	75	K	107	k
44	,	76	L	108	l
45	-	77	M	109	m
46	.	78	N	110	n
47	/	79	O	111	o
48	0	80	P	112	p
49	1	81	Q	113	q
50	2	82	R	114	r
51	3	83	S	115	s
52	4	84	T	116	t
53	5	85	U	117	u
54	6	86	V	118	v
55	7	87	W	119	w
56	8	88	X	120	x
57	9	89	Y	121	y
58	:	90	Z	122	z
59	;	91	[123	{
60	<	92	\	124	
61	=	93]	125	}
62	>	94	^	126	~
63	?	95	_		



Extensions du code ASCII

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Outre l'anglais, le jeu de caractères du code ASCII permet d'écrire des textes techniques de toute langue utilisant l'alphabet latin, si on accepte d'omettre les caractères accentués.

Cependant de telles omissions, ne sont pas acceptables pour des textes de nature littéraire. C'est pourquoi de nombreuses extensions du code ASCII sont très vite apparues.

Pour avoir les caractères accentués du français il faut utiliser l'ISO 8859-1 (parfois appelé LATIN-1) ou l'ISO 8859-15



Latin-1

Codage des nombres et des caractères

Bases de numération

Codage des nombres entiers

Codage des nombres réels

Codage des caractères

0	32	64	@	96	'	128	160	192	À	224	à		
1	33	!	65	A	97	a	129	161	í	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	¹	217	Ù	249	ù
26	58	:	90	Z	122	z	154	186	º	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



Code ASCII

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères

Les diverses extensions du code ASCII, dont les ISO 8859, présentent l'inconvénient d'être incompatibles entre-elles, et le plus souvent d'être spécialisées pour un jeu de caractères lié à une langue.

Comment alors coder dans un même document des textes rédigés avec des alphabets aussi divers que les alphabets latin, cyrillique, grec, arabe, ... ?



Unicode

Codage des
nombres et
des
caractères

Bases de
numération

Codage des
nombres
entiers

Codage des
nombres
réels

Codage des
caractères



L'Unicode a vu le jour au début des années 1990. Il recense actuellement environ 110000 caractères. Plusieurs codages des caractères Unicode existent. Le plus couramment utilisé, notamment pour les pages Web, est l'UTF-8 qui code chaque caractère sur 8, 16, 24 ou 32 bits (soit un, deux, trois ou quatre octets)

- ▶ Fonctions `chr()` et `ord()`
- ▶ comparaison de caractères