

Table ASCII

exemple Y correspond à x59 en hexadécimal c'est-à-dire 89 en décimal (d'abord le chiffre de la colonne). Le codage ASCII utilise un octet pour un caractère.

MSB \ LSB		0	1	2	3	4	5	6	7
		000	001	010	011	100	101	110	111
0	0000	NUL	DLE	SP	0	@	P	`	p
1	0001	SOH	DC1	!	1	A	Q	a	q
2	0010	STX	DC2	"	2	B	R	b	r
3	0011	ETX	DC3	#	3	C	S	c	s
4	0100	EOT	DC4	\$	4	D	T	d	t
5	0101	ENQ	NAK	%	5	E	U	e	u
6	0110	ACK	SYN	&	6	F	V	f	v
7	0111	BEL	ETB	'	7	G	W	g	w
8	1000	BS	CAN	(8	H	X	h	x
9	1001	HT	EM)	9	I	Y	i	y
A	1010	LF	SUB	*	:	J	Z	j	z
B	1011	VT	ESC	+	;	K	[k	}
C	1100	FF	FS	,	<	L	\	l	
D	1101	CR	GS	-	=	M]	m	{
E	1110	SO	RS	.	>	N	^	n	~
F	1111	SI	US	/	?	O	_	o	DEL

Début de la table UNICODE ...on retrouve exactement les mêmes codes, mais il faut lire d'abord le chiffre de la ligne.

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

Table ANSI

Pour avoir le code, on ajoute ici le nombre de la colonne et celui de la ligne.

Pour la partie gauche, on reconnaît la table ASCII. A droite les compléments, qui ne sont pas toujours compatibles avec l'UNICODE. On utilise toujours un octet !

Exemples : é codé 233 en ANSI

DEC	0	16	32	48	64	80	96	112	128	144	160	176	192	208	224	240
0			space	0	@	P	`	p	€	□	nbsp	°	À	Ð	à	ð
1			!	1	A	Q	a	q	□	´	ı	±	Á	Ñ	á	ñ
2			"	2	B	R	b	r	,	´	¸	²	Â	Ò	â	ò
3			#	3	C	S	c	s	f	“	£	³	Ã	Ó	ã	ó
4			\$	4	D	T	d	t	„	”	¤	´	Ä	Ô	ä	ô
5			%	5	E	U	e	u	…	•	¥	µ	Å	Ö	å	ö
6			&	6	F	V	f	v	†	-	ı	¶	Æ	Ø	æ	ø
7			'	7	G	W	g	w	‡	—	§	·	Ç	×	ç	÷
8			(8	H	X	h	x	^	ˆ	˘	,	È	Ø	è	ø
9	TAB)	9	I	Y	i	y	%o	™	©	¹	É	Ù	é	ù
10	LF		*	:	J	Z	j	z	Š	š	ª	º	Ê	Ú	ê	ú
11			+	;	K	[k	{	<	>	«	»	Ë	Û	ë	û
12			,	<	L	\	l		Œ	œ	¬	¼	Ì	Ü	ì	ü
13	CR		-	=	M]	m	}	□	□		½	Í	Ý	í	ý
14			.	>	N	^	n	~	Ž	ž	®	¾	Î	Þ	î	þ
15			/	?	O	_	o	□	□	ÿ	-	¿	Ï	ß	ï	ÿ

et en UNICODE : x00E9 ce qui donne bien 14*16+9 = 233

PDF : fr en v - d - m	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
008	PAD	HOP	BPH	NBH	IND	NEL	SSA	ESA	HTS	HTJ	VTS	PLD	PLU	RI	SS2	SS3
009	DCS	PU1	PU2	STS	CCH	MW	SPA	EPA	SOS	SGCI	SCI	CSI	ST	OSC	PM	APC
00A	NBSP	ı	¸	£	¤	¥	ı	§	ˆ	©	ª	«	¬	SHY	®	™
00B	°	±	²	³	´	µ	¶	·	,	¹	º	»	¼	½	¾	¿
00C	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
00D	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
00E	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
00F	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

mais le TM
a pour code ANSI 153

Symboles lettrés [modifier | modifier le code]

PDF : fr en v · d · m	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	
210	‰	‱	℄	°C	€	‰	‰	ε	Ϸ	°F	g	ℋ	ℋ	H	h	ħ	
211	ƒ	g	ℓ	ℓ	℥	N	№	©	ρ	P	Q	℞	℞	R	R	Ř	
212	SM	TEL	TM	Ÿ	Z	Ƴ	Ω	U	Ƴ	ı	K	Å	ß	€	e	e	
213	ε	ƒ	ƒ	ℳ	¢	κ	ı	λ	τ	ı	Ϡ	FAX	Ⓜ	γ	Γ	Π	
214	Σ	Ϡ	τ	J	λ	D	d	e	i	j	ℙ	℘	Ⓜ	¼	½	ƒ	Ⓜ

en UNICODE x 2122 (en hexadécimal) ce qui correspond à 8482 !

Attention, l'UNICODE n'est qu'un code numérique attribué à chaque caractère...
Si on voulait utiliser l'UNICODE pour encoder les caractères, en admettant qu'on se limite à ce qu'on appelle le plan 0, c'est-à-dire les caractères qui se codent avec 4 chiffres en hexadécimal (65536 caractères), il faudrait 16 bits, soit deux octets....
En ANSI, il ne faut qu'un seul octet....Pour un texte "normal", on occuperait donc presque deux fois plus de place en mémoire !

L'une des principales normes d'encodage des caractères est l'Utf-8.
Les principaux caractères (pour les anglosaxons !), ceux de la table ASCII, sont encodés sur un octet, mais d'autres sont encodés sur deux, trois ou quatre octets.
On voudrait décoder :

01000000 11100010 10000100 10100010 11000011 10111100

Evidemment, pour décoder, il faut savoir combien d'octets lire pour avoir un caractère !

Ce sont les bits de poids forts qui l'indiquent :

- un octet commençant par 0 se lit tout seul.
par exemple : b01000001 correspond à 65 qui est A (aussi bien en UNICODE que dans les tables ANSI et ASCII)
- un octet commençant par 11 se lit avec le suivant qui commence obligatoirement par 10 (et pour obtenir le code du caractère on enlève le 11 et le 10, on assemble les autres bits)
par exemple : 11000001 10100000 correspond à b00001100000 qui vaut 192, donc à xC0 en hexadécimal....donc à À

- un octet commençant par 111 se lit avec les deux suivants qui commenceront par 10
- un octet commençant par 1111 se lit avec les trois suivants qui commencent par 10

Exemple : décoder

01000000 11100010 10000100 10100010 11000011 10111100

On cherchera d'abord les codes UNICODE en hexadécimal des caractères et on utilisera internet pour retrouver les caractères correspondants.

On trouve 0040 2122 00fc ce qui donne @™ü