

O Código ASCII

Copyright ©

Este documento, intitulado 'O Código ASCII' e de autoria da CCM (<https://br.ccm.net/>), está disponível sob a licença Creative Commons.

Você pode copiar e/ou modificar o conteúdo desta página com base nas condições estipuladas pela licença.

Sumário

A codificação das informações.....	3
O que é o código ASCII	3
Tabela dos caracteres ASCII	3
O código EBCDIC.....	8
Código Unicode.....	8

A codificação das informações

O morse foi a primeira codificação que permitiu a comunicação remota. Foi Samuel F. B. Morse que o criou em 1844. Este código é composto de pontos e travessões (código binário, de certa forma). Ele permitiu efetuar comunicações muito mais rápidas do que com o sistema de correio da época nos Estados Unidos, o *Pony Express*. O intérprete era o homem, por isso era preciso um bom conhecimento do código. Vários códigos foram inventados posteriormente, incluindo o **código Baudot**, chamado de código Murray pelos ingleses. Em 10 de março de 1876, o Alexander Graham Bell criou o primeiro telefone, invenção revolucionária que permitiu fazer circular a informação vocal através de linhas metálicas. Estas linhas permitiram o desenvolvimento de máquinas para codificar e decodificar os caracteres pelo código Baudot. Nos anos 60, o **código ASCII** (American Standard Code for Information Interchange) foi adotado como código padrão. Ele permite a codificação de caracteres em 8 bits, dando a possibilidade de formação de 256 caracteres.

O que é o código ASCII

A memória do computador conserva todos os dados sob a forma digital. Não existe um método para armazenar diretamente os caracteres. Cada caractere possui o seu equivalente em código numérico, é o ASCII. O código ASCII básico representava os caracteres em 7 bits (ou seja, 128 caracteres possíveis, de 0 a 127). **Os códigos 0 a 31 não são caracteres** são chamados de caracteres de controle porque permitem fazer ações como carriage return (**CR** - retorno do carro em português) e bip sonoro (**BEL**), os **Os códigos 65 a 90 representam as maiúsculas** e os **Os códigos 97 a 122 representam as minúsculas**. Basta alterar o 6º bit para passar das maiúsculas para as minúsculas, isto é, acrescentar 32 ao código ASCII em base decimal.

Tabela dos caracteres ASCII

Caractere	Código ASCII	Código hexadecimal
NUL (<i>Nulo</i>)	0	00
SOH (<i>Start of heading</i>)	1	01
STX (<i>Start of text</i>)	2	02
ETX (<i>End of text</i>)	3	03

EOT (<i>End of transmission</i>)	4	04
ENQ (<i>Enquiry</i>)	5	05
ACK (<i>Acknowledge</i>)	6	06
BEL (<i>Bell</i>)	7	07
BS (<i>Backspace</i>)	8	08
TAB (Tabulação horizontal)	9	09
LF (<i>Line Feed</i> , salto de linha)	10	0A
VT (<i>Vertical tabulation</i> , tabulação vertical)	11	0B
FF (<i>Form feed</i>)	12	0C
CR (<i>Carriage return</i> , retorno do carro)	13	0D
SO (<i>Shift out</i>)	14	0E
SI (<i>Shift in</i>)	15	0F
DLE (<i>Data link escape</i>)	16	10
DC1 (<i>Device control 1</i>)	17	11
DC2 (<i>Device control 2</i>)	18	12
DC3 (<i>Device control 3</i>)	19	13
DC4 (<i>Device control 4</i>)	20	14
NAK (<i>Negative acknowledgement</i>)	21	15
SYN (<i>Synchronous idle</i>)	22	16
ETB (<i>End of transmission block</i> , fim de bloco de transmissão)	23	17
CAN (<i>Cancel</i> , cancelar)	24	18
EM (<i>End of medium</i> , fim do meio)	25	19
SUB (<i>Substitute</i> , substituto)	26	1A
ESC (<i>Escape</i> , caractere ESC)	27	1B
FS (<i>File separator</i> , separador de arquivo)	28	1C
GS (<i>Group separator</i> , separador de grupo)	29	1D
RS (<i>Record separator</i> , separador de registro)	30	1E
US (<i>Unit separator</i> , separador de registro)	31	1F
SP (<i>Space</i> , espaço)	32	20
!	33	21
"	34	22
#	35	23
\$	36	24

%	37	25
&	38	26
'	39	27
(40	28
)	41	29
*	42	2A
+	43	2B
,	44	2C
-	45	2D
.	46	2E
/	47	2F
0	48	30
1	49	31
2	50	32
3	51	33
4	52	34
5	53	35
6	54	36
7	55	37
8	56	38
9	57	39
:	58	3A
;	59	3B
<	60	3C
=	61	3D
>	62	3E
?	63	3F
@	64	40
A	65	41
B	66	42
C	67	43
D	68	44
E	69	45

F	70	46
G	71	47
H	72	48
I	73	49
J	74	4A
K	75	4B
L	76	4C
M	77	4D
N	78	4E
O	79	4F
P	80	50
Q	81	51
R	82	52
S	83	53
T	84	54
U	85	55
V	86	56
W	87	57
X	88	58
Y	89	59
Z	90	5A
[91	5B
	92	5C
]	93	5D
^	94	5E
_	95	5F
'	96	60
a	97	61
b	98	62
c	99	63
d	100	64
e	101	65
f	102	66

g	103	67
h	104	68
i	105	69
j	106	6A
k	107	6B
l	108	6C
m	109	6D
n	110	6E
o	111	6F
p	112	70
q	113	71
r	114	72
s	115	73
t	116	74
u	117	75
v	118	76
w	119	77
x	120	78
y	121	79
z	122	7A
{	123	7B
	124	7C
}	125	7D
~	126	7E
Toque de supressão	127	7F

Tabela de caracteres ASCII estendido

O código **ASCII** foi concebido para a língua inglesa, por isso ele não contém caracteres acentuados, nem caracteres específicos de outro idioma. Para codificar este tipo de caractere é necessário recorrer a outro código. O código ASCII foi então estendido para 8 bits (um byte) para poder codificar mais caracteres (fala-se, aliás, de código ASCII estendido). Este código atribui os valores de 0 a 255 às letras maiúsculas e minúsculas, aos números, às marcas de pontuação e aos outros símbolos (caracteres acentuados no caso do código iso-latin1). Ele não é único e depende muito da plataforma utilizada. Os dois jogos de caracteres ASCII estendidos mais frequentemente utilizados são o código ASCII estendido **OEM**, ou seja, aquele que equipava as primeiras máquinas de tipo PC da IBM:

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
8	ç	ü	é	â	ä	à	æ	ç	ê	ë	è	ï	î	ì	ñ	ø
9	é	æ	œ	ô	ö	ò	û	ù	ÿ	ö	ü	ç	£	¥	℞	ƒ
A	á	í	ó	ú	ñ	Ñ	º	º	¿	¿	½	¾	¿	«	»	
B	▨	▩	▪													
C	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞
D	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞	⌞
E	α	β	Γ	Π	Σ	σ	μ	τ	ϑ	θ	Ω	δ	ω	ϕ	€	π
F	≡	±	≥	≤	∫	J	÷	≈	°	·	·	√	π	²	■	

Também há o código ASCII estendido **ANSI**, utilizado pelos sistemas operacionais recentes:

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
8	□	□	,	f	„	...	†	‡	^	‰	Š	<	œ	□	□	□
9	□	‚	‚	“	”	•	—	—	˜	™	Š	>	œ	□	□	Ÿ
A		ı	ç	£	¤	¥	¦	§	¨	©	ª	«	¬	­	®	¯
B	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
C	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
D	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
E	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
F	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

O código EBCDIC

O código **EBCDIC** (Extended Binary - Coded Decimal Interchange Code), desenvolvido pela IBM, permite codificar caracteres em 8 bits. Embora muito utilizado nas máquinas IBM, não teve o sucesso do código ASCII.

Código Unicode

O sistema **Unicode** é um sistema de codificação dos caracteres em 16 bits inventado em 1991. O sistema Unicode permite representar qualquer caractere por um código em 16 bits, independentemente de qualquer sistema operacional ou linguagem de programação. Ele reúne, assim, quase todos os alfabetos existentes (árabe, armênio, cirílico, grego, hebreu, latino, etc.) e é compatível com o código ASCII. O conjunto dos códigos Unicode está disponível no site <http://www.unicode.org/charts/>.