

Miscellaneous Character Sets

This appendix is not included in the printed version of this book, and is instead available as a downloadable and printable PDF file. As new material becomes available, the PDF file will be updated accordingly.*

The material currently included is directly excerpted from the first edition, and an updated appendix for the second edition will be made available as that material becomes available.

* <http://examples.oreilly.com/9780596514471/cjkvip2e-appM.pdf>

S

Single-Byte Code Tables

This appendix provides code tables that illustrate the characters in a variety of single-byte character sets, indexed by hexadecimal values. Control characters, when defined, appear as visual tags.

Non-CJKV Code Tables

This section provides code tables for non-CJKV character sets, such as ASCII, EBCDIC, and ISO 8859-1:1998. All of these character sets, as described in the beginning sections of Chapter 3, *Character Set Standards*, form the foundation for locale-specific instances.

ASCII Code Table

The following is an ASCII code table, including the C0 control code region (0x00 through 0x1F and 0x7F) expressed as visual tags:

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

EBCDIC Code Table

The following is an EBCDIC code table, including control characters expressed as visual tags:

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	<NUL>	<SOH>	<STX>	<ETX>	<SEL>	<HT>	<RNL>		<GE>	<SPS>	<RPT>	<VT>	<FF>	<CR>	<SO>	<SI>
1	<DLE>	<DC1>	<DC2>	<DC3>	<RES>	<NL>	<BS>	<POC>	<CAN>		<UBS>	<CU1>	<IFS>	<IGS>	<IRS>	<IUS>
2	<DS>	<SOS>	<FS>	<WUS>	<BYP>	<LF>	<ETB>	<ESC>	<SA>	<SFE>	<SM>	<CSP>	<MFA>	<ENQ>	<ACK>	<BEL>
3			<SYN>	<IR>	<PP>	<TRN>	<NBS>	<EOT>	<SBS>	<IT>	<RFF>	<CU3>	<DC4>	<NAK>		<SUB>
4	<SP>										[.	<	(+	!
5	&]	\$	*)	;	^
6	-	/										,	%	_	>	?
7										`	:	#	@	'	=	"
8		a	b	c	d	e	f	g	h	i						
9		j	k	l	m	n	o	p	q	r						
A		~	s	t	u	v	w	x	y	z						
B																
C	{	A	B	C	D	E	F	G	H	I						
D	}	J	K	L	M	N	O	P	Q	R						
E	\		S	T	U	V	W	X	Y	Z						
F	0	1	2	3	4	5	6	7	8	9						

ISO 8859-1:1998 Code Table

The table at the top of the following page illustrates the complete ISO 8859-1:1998 character set and its encoding. Note how the region up until 0x7F is identical to ASCII, which is a common property for the remaining parts of ISO 8859. The control codes <NBSP> and <SHY> stand for “non-breaking space” and “soft hyphen,” respectively.

Chinese Code Tables

This section provides Chinese single-byte code tables, specifically for GB-Roman and CNS-Roman.