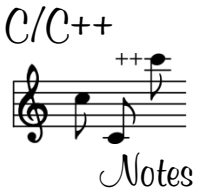
Appendix B



The ASCII Character Codes

NOTE B.1

The ASCII Character Codes

Control Characters

DEC	OCT	HEX	CHR	DEC	OCT	HEX	CHR	DEC	OCT	HEX	CHR	DEC	OCT	HEX	CHR
000	000	00	NUL	032	040	20	<sp>	064	100	40	@	096	140	60	`
001	001	01	SOH	033	041	21	!	065	101	41	A	097	141	61	a
002	002	02	STX	034	042	22	"	066	102	42	B	098	142	62	b
003	003	03	ETX	035	043	23	#	067	103	43	C	099	143	63	c
004	004	04	EOT	036	044	24	$	068	104	44	D	100	144	64	d
005	005	05	ENQ	037	045	25	%	069	105	45	E	101	145	65	e
006	006	06	ACK	038	046	26	&	070	106	46	F	102	146	66	f
007	007	07	BEL	039	047	27	'	071	107	47	G	103	147	67	g
008	010	08	BS	040	050	28	(	072	110	48	H	104	150	68	h
009	011	09	HT	041	051	29	)	073	111	49	I	105	151	69	i
010	012	0a	LF	042	052	2a	*	074	112	4a	J	106	152	6a	j
011	013	0b	VT	043	053	2b	+	075	113	4b	K	107	153	6b	k
012	014	0c	FF	044	054	2c	,	076	114	4c	L	108	154	6c	l
013	015	0d	CR	045	055	2d	-	077	115	4d	M	109	155	6d	m
014	016	0e	SO	046	056	2e	.	078	116	4e	N	110	156	6e	n
015	017	0f	SI	047	057	2f	/	079	117	4f	O	111	157	6f	o
016	020	10	DLE	048	060	30	0	080	120	50	P	112	160	70	p
017	021	11	DC1	049	061	31	1	081	121	51	Q	113	161	71	q
018	022	12	DC2	050	062	32	2	082	122	52	R	114	162	72	r
019	023	13	DC3	051	063	33	3	083	123	53	S	115	163	73	s
020	024	14	DC4	052	064	34	4	084	124	54	T	116	164	74	t
021	025	15	NAK	053	065	35	5	085	125	55	U	117	165	75	u
022	026	16	SYN	054	066	36	6	086	126	56	V	118	166	76	v
023	027	17	ETB	055	067	37	7	087	127	57	W	119	167	77	w
024	030	18	CAN	056	070	38	8	088	130	58	X	120	170	78	x
025	031	19	EM	057	071	39	9	089	131	59	Y	121	171	79	y
026	032	1a	SUB	058	072	3a	:	090	132	5a	Z	122	172	7a	z
027	033	1b	ESC	059	073	3b	;	091	133	5b	[	123	173	7b	{
028	034	1c	FS	060	074	3c	<	092	134	5c	\	124	174	7c	|
029	035	1d	GS	061	075	3d	=	093	135	5d	]	125	175	7d	}
030	036	1e	RS	062	076	3e	>	094	136	5e	^	126	176	7e	~
031	037	1f	US	063	077	3f	?	095	137	5f	_	127	177	7f	del

The values of all successive members of the numeric and alphabetic character sets are sequential. Thus, there is a fixed difference between corresponding members of the upper and lower case character sets and a similar fixed difference between the numeric and alphabetic character sets. The first 32_{10} characters are called "control characters" because they have been historically used for message control rather than for displaying actual printable characters. Such message control can range anywhere from moving the screen cursor to the next line to implementing complex data communication protocols. Two different names are commonly associated with each control character:

A. The first is based upon how the character is commonly used in data communication protocols. In this case the character's name appears as an abbreviation in the table. For example, BS stands for "backspace", LF stands for "line feed", SYN stands for "synchronize", ESC stands for "escape", etc.

B. The second is based upon the character's physical position in the table itself. That is, BS, LF, SYN, and ESC from the previous paragraph become "control-H", "control-J", "control-V", and "control-left-bracket". These are often abbreviated ^H, ^J, ^V, and ^[for ease of documentation. The correct name for any control character is easily determined by simply adding 64_{10} (same as 40_{16} and 100_8) to the value of that character and then putting the word "control-" in front of the resulting character.

WHAT IS UNICODE?

Unicode is an internationally standardized encoding scheme intended to provide a unique platform-independent numeric representation for every approved character in every approved written language in the world. It is compatible with ASCII in that the first 128 characters of each have the same numeric values (ASCII only has 128 characters).

