

ASCII は文字コードの基本



GN0001-00 / 2023 年 06 月 25 日 / こがねさん(著)



<https://www.kumikomist.com/>

■目次

1. ASCII とは.....	2	2. JIS X 0201	4
1.1. ASCII 表	2	3. おまけ	5
1.2. 制御文字.....	3	3.1. キャラクターLCD	5
1.3. ASCII の特徴.....	4		

■文書内の記号

	取り扱いにおける禁止事項（してはいけないこと）を示しています。
	取扱における指示事項（必ずしなければいけないこと）を示しています。

	取り扱いにおける注記事項を示しています。
	取り扱いにおけるポイントを示しています。

1. ASCII とは

ASCII (American Standard Code for Information Interchange) は、7 桁の 2 進数で表すことのできる整数の値にラテン文字、数字、記号、制御文字を割り当てた文字コードです。なお「ASCII コード」と表記されることもあります。また、「C = コード」なのでコードが 2 重に掛かり正しくありません。

ASCII は 7 ビットコードで、英語以外の言語には対応していません。そこで各国は 8 ビットコードに拡張して 80h~FFh に独自の文字や記号を割り当てました。「2. JIS X 0201」で紹介する日本の「JIS X 0201」もその一つです。この結果各国間で文字コードに互換性がなくなり、違う国のコンピューターで見ると文字化けしてしまうという負の時代が続きました。

この問題を解決すべく生まれた国際的な統一文字コード UTF-8 も ASCII には配慮しており、0~127 番は ASCII そのものとなっています。

1.1. ASCII 表

ASCII は大きく分けると 00h~1Fh と 7Fh を含めた制御文字と、その他の印字可能文字・記号に分かれています。

表 1 ASCII 表

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

1.2. 制御文字

制御文字は表示するための文字や記号ではありません。モニターやプリンターなどの機器を制御するために用いられます。STXやETX、CRやLFはRS-232Cなどの通信電文でもよく用いられています。

表 2 ASCII 制御文字

16 進数	略語	語源	意味
00	NUL	Null	ヌル文字
01	SOH	Start of Heading	ヘッディング開始
02	STX	Start of Text	テキスト開始
03	ETX	End of Text	テキスト終了
04	EOT	End of Transmission	伝送終了
05	ENQ	Enquiry	問い合わせ
06	ACK	Acknowledgement	肯定応答
07	BEL	Bell	ベル
08	BS	Back Space	後退
09	HT	Horizontal Tabulation	水平タブ
0A	LF	Line Feed	改行
0B	VT	Vertical Tabulation	垂直タブ
0C	FF	Form Feed	書式送り
0D	CR	Carriage Return	復帰
0E	SO	Shift Out	シフトアウト
0F	SI	Shift In	シフトイン
10	DLE	Data Link Escape	伝送制御拡張
11	DC1	Device Control 1	装置制御 1、XON
12	DC2	Device Control 2	装置制御 2
13	DC3	Device Control 3	装置制御 3、XOFF
14	DC4	Device Control 4	装置制御 4
15	NAK	Negative Acknowledgement	否定応答
16	SYN	Synchronous Idle	同期信号
17	ETB	End of Transmission Block	伝送ブロック終結
18	CAN	Cancel	取消
19	EM	End of Medium	媒体終端
1A	SUB	Substitute	置換
1B	ESC	Escape	エスケープ
1C	FS	File Separator	ファイル分離標識
1D	GS	Group Separator	グループ分離標識
1E	RS	Record Separator	レコード分離標識
1F	US	Unit Separator	ユニット分離標識
7F	DEL	Delete	抹消

1.3. ASCII の特徴

- (1) 大文字の ASCII 値の 5 ビット目を 1 にすると、小文字になります。
例: 'A' ⇒ 41h ⇒ 61h ⇒ 'a'
- (2) 数字の上位 4 ビットを 0 にすると、対応する値が得られます。
例: '0' ⇒ 30h ⇒ 00h ⇒ 0

2. JIS X 0201

JIS X 0201 は「7 ビット及び 8 ビットの情報交換用符号化文字集合」で、ASCII のラテン文字にカタカナを追加した文字コードです。空白の部分は未定義です。



- 5Ch と 7Eh は ASCII と文字が異なるので注意してください。

表 3 JIS X 0201

	2-	3-	4-	5-	6-	7-	A-	B-	C-	D-	E-	F-
-0	/	0	@	P	`	p	/	ー	タ	ミ		
-1	!	1	A	Q	a	q	。	ア	チ	ム		
-2	"	2	B	R	b	r	「	イ	ツ	メ		
-3	#	3	C	S	c	s	」	ウ	テ	モ		
-4	\$	4	D	T	d	t	、	エ	ト	ヤ		
-5	%	5	E	U	e	u	・	オ	ナ	ユ		
-6	&	6	F	V	f	v	ヲ	カ	ニ	ヨ		
-7	'	7	G	W	g	w	ァ	キ	ヌ	ラ		
-8	(8	H	X	h	x	ィ	ク	ネ	リ		
-9)	9	I	Y	i	y	ゥ	ケ	ノ	ル		
-A	*	:	J	Z	j	z	ェ	コ	ハ	レ		
-B	+	;	K	[k	{	ォ	サ	ヒ	ロ		
-C	,	<	L	¥	l		ャ	シ	フ	ワ		
-D	-	=	M]	m	}	ュ	ス	ヘ	ン		
-E	.	>	N	^	n	ー	ョ	セ	ホ	ゞ		
-F	/	?	O	_	o	/	ッ	ソ	マ	。		/

3. おまけ

3.1. キャラクターLCD

組み込みシステムでよく用いられているキャラクターLCD というものがあります。これは横 5 ドット×縦 8 ドットで文字を表すもので、JIS X 0201 と殆ど同じ並びとなっています。さらに E0h~FFh の空きエリアに、特殊な記号などを割り当てられてるのが一般的です。メーカーによって記号が一部変わりますが、概ね下図のようになっています。

	2-	3-	4-	5-	6-	7-	A-	B-	C-	D-	E-	F-
-0		0	0	P	`	P		—	9	3	0	P
-1	!	1	A	Q	a	q		7	7	4	3	Q
-2	"	2	B	R	b	r	「	イ	ウ	×	0	0
-3	#	3	C	S	c	s	」	ウ	T	0	0	0
-4	\$	4	D	T	d	t	,	工	ト	カ	ウ	0
-5	%	5	E	U	e	u	.	オ	ナ	コ	0	0
-6	&	6	F	V	f	v	ヲ	カ	ニ	ヨ	0	Σ
-7	'	7	G	W	g	w	7	+	ヌ	ラ	Q	π
-8	(8	H	X	h	x	イ	ウ	ホ	リ	ル	又
-9)	9	I	Y	i	y	ウ	ウ	ツ	ル	”	U
-A	*	:	J	Z	j	z	エ	コ	ハ	レ	i	7
-B	+	:	K	C	k	c	オ	ウ	ヒ	ロ	×	π
-C	,	<	L	*	l	l	カ	シ	フ	ワ	0	π
-D	—	=	M	I	m	i	ユ	ズ	ハ	ソ	ト	÷
-E	.	>	N	^	n	^	ヨ	セ	ホ	”	π	
-F	/	?	O	_	o	_	ッ	リ	マ	”	0	■

図 1 キャラクターLCD のドット例

ドット絵だとわかりにくいので、文字に置き換えたものが下記表になります。

表 4 キャラクター-LCD の文字例

	2-	3-	4-	5-	6-	7-	A-	B-	C-	D-	E-	F-
-0		0	@	P	`	p		一	夕	ミ	α	p
-1	!	1	A	Q	a	q	。	ア	チ	ム	ä	q
-2	"	2	B	R	b	r	「	イ	ツ	メ	β	θ
-3	#	3	C	S	c	s	」	ウ	テ	モ	ε	∞
-4	\$	4	D	T	d	t	、	エ	ト	ヤ	μ	Ω
-5	%	5	E	U	e	u	・	オ	ナ	ユ	σ	ü
-6	&	6	F	V	f	v	ヲ	カ	ニ	ヨ	ρ	Σ
-7	'	7	G	W	g	w	ア	キ	ヌ	ラ	g	π
-8	(8	H	X	h	x	イ	ク	ネ	リ	√	̄
-9)	9	I	Y	i	y	ウ	ケ	ノ	ル	⁻¹	y
-A	*	:	J	Z	j	z	エ	コ	ハ	レ	j	千
-B	+	;	K	[k	{	オ	サ	ヒ	口	x	万
-C	,	<	L	¥	l		ヤ	シ	フ	ワ	¢	円
-D	-	=	M]	m	}	ユ	ス	ヘ	ン	£	÷
-E	.	>	N	^	n	→	ヨ	セ	ホ	ˆ	ñ	
-F	/	?	O	_	o	←	ツ	ソ	マ	°	ö	■

表 4 の E0h~FFh に割り当てられている特殊な記号が何を表しているのか、思うところを記しておきます。

- (1) 「g」「j」「p」「q」「y」はディセンダー付きデザインです。このディセンダー付き文字ですが、実は縦 10 ドットでデザインされています(下図赤点部)。このため縦 8 ドットの LCD では下 2 ドット分が描画されず、切れた形になっています。「β」や「μ」も同様です。なお ASCII 値の 7 ビット目を 1 にした位置に配置されています。

	-0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-A
E-	α	ä	β	ε	μ	σ	ρ	q	r	'	j
F-	p	q	θ	ω	Ω	ü	Σ	π	×	u	千

- (2) 「ä」「ñ」「ö」「ü」はアクセント付きの文字です。ラテン文字ではない何か。ASCII 値の 7 ビット目を 1 にした位置に配置されています。
- (3) 「¢」はセント、「£」はポンドだと思います。
- (4) 「x」は何も想像がつかない記号です。「x (かける)」の上付きのように見えますが???