A.4.13 SMS/CBS 7-bit Urdu (Eastern Arabic script) Alphabet (0x13)

A.4.13.1 SMS/CBS 7-bit Urdu (Eastern Arabic script) Alphabet (0x13) base table

				b7	0	0	0	0	1	1	1	1
				b6	0	0	1	1	0	0	1	1
				b5	0	1	0	1	0	1	0	1
b4	**b3**	**b2**	**b1**		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	0	0	0	**0**	ا 0627	آ 0622	**SP** 0020	0 0030	ص 0635	ں 06BA	�̊ 0652	p 0070
0	0	0	1	**1**	ب 0628	_ 005F	ڏ 068F	1 0031	ض 0636	ٹ 06BB	a 0061	q 0071
0	0	1	0	**2**	ٻ 067B	ح 062D	ڍ 068D	2 0032	ط 0637	ڼ 06BC	b 0062	r 0072
0	0	1	1	**3**	پ 067E	خ 062E	# 0023	3 0033	ظ 0638	و 0648	c 0063	s 0073
0	1	0	0	**4**	ڀ 0680	ج 062C	ذ 0630	4 0034	ع 0639	ۄ 06C4	d 0064	t 0074
0	1	0	1	**5**	ت 062A	ځ 0681	ر 0631	5 0035	غ 063A	ؤ 0624	e 0065	u 0075
0	1	1	0	**6**	ٿ 067C	ڃ 0683	ز 0632	6 0036	ق 0642	ۆ 06C6	f 0066	v 0076
0	1	1	1	**7**	ٽ 067D	ڄ 0684	ڑ 0691	7 0037	ک 06A9	ء 0621	g 0067	w 0077
S	0	0	0	**8**	ٹ 0679	څ 0685	ژ 0698	8 0038	ڪ 06AA	ى 06CC	h 0068	x 0078
1	0	0	1	**9**	ث 062B	چ 0686	ڙ 0699	9 0039	ګ 06AB	ې 06D0	i 0069	y 0079
1	0	1	0	**A**	**LF** 000A	د 062F	* 002A	: 003A	گ 06AF	ے 06D2	j 006A	z 007A
1	0	1	1	**B**	ٺ 067A	**SS2** ----	+ 002B	; 003B	ڳ 06B3	◌َ 064E	k 006B	، 060C
1	1	0	0	**C**	ٿ 067F	ڌ 068C	, 002C	س 0633	ڱ 06B1	◌ِ 0650	l 006C	۔ 06D4
1	1	0	1	**D**	**CR** 000D	ڈ 0688	- 002D	ښ 069A	ل 0644	◌ُ 064F	m 006D	ـ 0640
1	1	1	0	**E**	ف 0641	ډ 0689	. 002E	ش 0634	م 0645	◌ّ 0651	n 006E	**LRM** 200E
1	1	1	1	**F**	ڦ 06A6	ڊ 068A	/ 002F	? 003F	ن 0646	◌ٓ 0653	o 006F	**RLM** 200F

The Arabic letters are displayed right-to-left, and in addition, Arabic letters are cursively joined. *Simplified* Unicode bidi algorithm for SMS/CBS: Bidi AL must be handled as bidi R, except Arabic punctuation which is handled as ON; all bidi EN and AN handled as L; all ET, ES, CS handled as ON, and 'paragraph direction' is always LTR, there are no bidi controls that span a text range. Bidi mirroring is done *only* by using the data in BidiMirroring.txt. Use a Nastaʿlīq style font for the display of the Arabic characters here.

LF: LINE FEED; moves to the next line (with implied carriage return).

CR: CARRIAGE RETURN; CR is not used as CR but is used as a filler after the actual SMS/CBS message text. CR should not occur inside a message text, but if it does, it must be converted to LF. Note though that <SS2,CR> is actually CSI (se next table), and that "CR" is not a filler, nor LF.

SS2: SINGLE SHIFT TWO; This code shifts the next 7-bit code unit to refer to an extension of this table (subclause A.4.13.2). When splitting a message text into submessages, there must be no cut right after an SS2.

ZWNJ: Use MIDDLE DOT (U+00B7, 0x1B3F) (visible) to get non-join shape(s) instead of ZWNJ (invisible).

ZWJ: Use ARABIC TATWEEL (U+0640, 0x7C) (visible) to get a join shape instead of ZWJ (invisible).

LRM,RLM: These should only be used to get the correct 'bidi directionality' of initially 'bidi neutral' symbols in case of *bidi directional change* around the symbol, to get the correct (readable!) display order. While zero-width, these should have visible glyphs; like a narrow bar with a 'flag' indicating direction.

A.4.13.2 SMS/CBS 7-bit Urdu (Eastern Arabic script) Alphabet (0x13) extension (SS2) table

b7					0	0	0	0	1	1	1	1
b6					0	0	1	1	0	0	1	1
b5					0	1	0	1	0	1	0	1
b4	**b3**	**b2**	**b1**	**1B**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	0	0	0	0	@ 0040	< 003C	۴ 06F4	، 060D	| 007C	P 0050	¼ 00BC	⚠ 26A0
0	0	0	1	1	£ 00A3	= 003D	۵ 06F5	ٔ 0654	A 0041	Q 0051	½ 00BD	∞ 221E
0	0	1	0	2	$ 0024	> 003E	۶ 06F6	ٕ 0655	B 0042	R 0052	¾ 00BE	² 00B2
0	0	1	1	3	¥ 00A5	₹ 20B9	۷ 06F7	أ 0623	C 0043	S 0053	" 201C	³ 00B3
0	1	0	0	4	§ 00A7	^ 005E	۸ 06F8	ٲ 0672	D 0044	T 0054	" 201D	ڇ 0687
0	1	0	1	5	" 0022	٫ 066B	۹ 06F9	ٳ 0673	E 0045	U 0055	¤ 00A4	ه 0647
0	1	1	0	6	€ 20AC	٬ 066C		ي 064A	F 0046	V 0056	Ω 03A9	ہ 06C1
0	1	1	1	7	% 0025	(0028	ٟ 065F	ؠ 0620	G 0047	W 0057	µ 03BC	ۂ 06C2
1	0	0	0	8	& 0026	) 0029	{ 007B	ئ 0626	H 0048	X 0058	° 00B0	ھ 06BE
1	0	0	1	9	' 0027	। 0964	} 007D	ى 06CD	I 0049	Y 0059	× 00D7	ً 064B
1	0	1	0	A	**FF** 000C	॥ 0965	۱⁄۴ 06F1,2044,06F4	≈ 2248	J 004A	Z 005A	÷ 00F7	ٍ 064D
1	0	1	1	B	− 2212	**SS3** ----	۱⁄۲ 06F1,2044,06F2	≠ 2260	K 004B	؛ 061B	± 00B1	ٌ 064C
1	1	0	0	C	! 0021	۰ 06F0	۳⁄۴ 06F3,2044,06F4	[005B	L 004C	؟ 061F	‰ 2030	ٗ 0657
1	1	0	1	D	**CSI** 009B	۱ 06F1	⩽ 2A7D	~ 007E	M 004D	٪ 066A	« 00AB	٘ 0658
1	1	1	0	E	ږ 0696	۲ 06F2	⩾ 2A7E	] 005D	N 004E	؉ 0609	» 00BB	ٖ 0656
1	1	1	1	F	ړ 0693	۳ 06F3	\ 005C	· 00B7	O 004F	ݐ 0690	' 2019	ٰ 0670

In the event that an MS receives a code where a symbol is not represented in the above table then the MS shall display the REPLACEMENT CHARACTER (U+FFFD).

FF: FORM FEED; This code is defined as a Page Break character (with implied carriage return). Any mobile station which does not have pagination shall treat is as a LF (with implied carriage return).

CSI: CONTROL SEQUENCE INTRODUCER; enables decimal character references, enables styling. An MS which does not support CSI shall convert it to REPLACEMENT CHARACTER.

SS3: SINGLE SHIFT THREE; This code is reserved for the extension to another extension table. When splitting a message text into submessages, there must be no cut right after an SS3.

U+2044 FRACTION SLASH; the sequence may need to be isolated from adjacent digits by '+' sign(s)..