

A.4.14 SMS/CBS 7-bit Bengali/Bangla Alphabet (0x14)

A.4.14.1 SMS/CBS 7-bit Bengali/Bangla Alphabet (0x14) 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	০	ঁ 0981	ং 0990	ঃ SP 0020	০ 0030	ক 09AC	ঠ 09BE		p 0070	
0	0	0	1	1	১	ং 0982	ং 005F	ং 09A4	১ 0031	ং 09AD	ং 09BF	a 0061	q 0071	
0	0	1	0	2	২	ং 0983		ঁ 099F	২ 0032	ং 09AE	ঁ 09C0	b 0062	r 0072	
0	0	1	1	3	৩	অ 0985	ও 0993	# 0023	৩ 0033	ঘ 09AF	ঁ 09C1	c 0063	s 0073	
0	1	0	0	4	৪	আ 0986	ও 0994	থ 09A5	৪ 0034	঱ 09B0	ঁ 09C2	d 0064	t 0074	
0	1	0	1	5	৫	ই 0987	ক 0995	ঁ 09A0	৫ 0035		ঁ 09C3	e 0065	u 0075	
0	1	1	0	6	৬	জ 0988	খ 0996	ঁ 09A6	৬ 0036	ল 09B2	ঁ 09E2	f 0066	v 0076	
0	1	1	1	7	৭	উ 0989	গ 0997	ড 09A1	৭ 0037			g 0067	w 0077	
1	0	0	0	8	৮	উ 098A	ঘ 0998	ধ 09A7	৮ 0038		ঁ 09C7	h 0068	x 0078	
1	0	0	1	9	৯	ঁ 098B	ঁ 0999	ঁ 09A2	৯ 0039			i 0069	y 0079	
1	0	1	0	A	A	LF 000A	ঁ 099A	*	:	স 09B8	ঁ 09C8	j 006A	z 007A	
1	0	1	1	B	B	ঁ 098C	SS2 ----	+	;	ঁ 09B7	ঁ 09D7	k 006B		
1	1	0	0	C			ছ 099B	,	ণ 09A3	শ 09B6		l 006C		
1	1	0	1	D	D	CR 000D	জ 099C	ন 09A8	প 09AA	হ 09B9	ঁ 09CB	m 006D		
1	1	1	0	E	E	ঁ 098F	ঁ 099D	.	ফ 09AB	ঁ 09BC	ঁ 09CC	n 006E	□ 09FD	
1	1	1	1	F			ঁ 099E	/	?		ঁ 09CD	ঁ 006F		

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).

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 (see 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.4.14.2).

When splitting a message text into submessages, there must be no cut right after an SS2.

09CD: BENGALI SIGN VIRAMA: This is a control character for forming conjuncts. If the display font does not support the requested conjunct, U+09CD is a combining character with nominal glyph.

A.4.14.2 SMS/CBS 7-bit Bengali/Bangla Alphabet (0x14) 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
0	0	0	0	0	@ 0040	< 003C	8 09EA	ঝ 09E0	। 007C
0	0	0	1	1	£ 00A3	= 003D	৷ 09EB	৷ 09E1	Q 0041
0	0	1	0	2	\$ 0024	> 003E	ঢ 09EC	ঢ 09C4	R 0042
0	0	1	1	3	¥ 00A5	₹ 20B9	ঠ 09ED	ঠ 09E3	S 0043
0	1	0	0	4	§ 00A7	^ 005E	৮ 09EE		T 0044
0	1	0	1	5	" 0022	৯ 09F3	৯ 09EF		” 0045
0	1	1	0	6	€ 20AC	□ 0980			V 0046
0	1	1	1	7	% 0025	(0028			W 0047
1	0	0	0	8	& 0026) 0029	{ 007B	H 0048	X 0058
1	0	0	1	9	' 0027	I 0964	} 007D	I 0049	Y 0059
1	0	1	0	A	FF 000C	॥ 0965	ঘ8 09E7,2044,09EA	≈ 2248	Z 004A
1	0	1	1	B	- 2212	SS3 ----	ঘৰ 09E7,2044,09E8	≠ 2260	K 004B
1	1	0	0	C	! 0021	০ 09E6	ঘৰ 09E9,2044,09EA	[005B	L 004C
1	1	0	1	D	CSI 009B	ঢ 09E7	ঢ 2A7D	~ 007E	M 004D
1	1	1	0	E	- 002D	ঢ 09E8	ঢ 2A7E] 005D	N 004E
1	1	1	1	F		ঢ 09E9	ঢ 005C	.	O 004F

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 it 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 ZWNJ.