

## A.4.21 SMS/CBS 7-bit Sinhala Alphabet (0x21)

### A.4.21.1 SMS/CBS 7-bit Sinhala Alphabet (0x21) base table

					<b>b7</b>	0	0	0	0	1	1	1	1
					<b>b6</b>	0	0	1	1	0	0	1	1
					<b>b5</b>	0	1	0	1	0	1	0	1
<b>b4</b>	<b>b3</b>	<b>b2</b>	<b>b1</b>			0	1	2	3	4	5	6	7
0	0	0	0	0	අ	ං	ඃ	ඇ	ඉ	උ	ඌ	ඍ	ඏ
0	0	0	1	1	ං	ං	ං	ං	ං	ං	ං	ං	ං
0	0	1	0	2	ං	ං	ං	ං	ං	ං	ං	ං	ං
0	0	1	1	3	ං	ං	ං	ං	ං	ං	ං	ං	ං
0	1	0	0	4	ං	ං	ං	ං	ං	ං	ං	ං	ං
0	1	0	1	5	ං	ං	ං	ං	ං	ං	ං	ං	ං
0	1	1	0	6	ං	ං	ං	ං	ං	ං	ං	ං	ං
0	1	1	1	7	ං	ං	ං	ං	ං	ං	ං	ං	ං
1	0	0	0	8	ං	ං	ං	ං	ං	ං	ං	ං	ං
1	0	0	1	9	ං	ං	ං	ං	ං	ං	ං	ං	ං
1	0	1	0	ං	LF 000A	ං	*	:	ං	ං	ං	ං	ං
1	0	1	1	ං	SS2 ----	+	;	ං	ං	ං	ං	ං	ං
1	1	0	0	ං		ං	,	ං	ං	ං	ං	ං	ං
1	1	0	1	ං	CR 000D	ං	ං	ං	ං	ං	ං	ං	ං
1	1	1	0	ං	ං	ං	.	ං	ං	ං	ං	ං	ං
1	1	1	1	ං	ං	ං	/	?	ං	ං	ං	ං	ං

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.4.21.2).

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

U+0DCA: SINHALA SIGN AL-LAKUNA: Virama; This is a control character for forming conjuncts. If the display font does not support the requested conjunct, U+0DCA is a combining character with nominal glyph.

#### A.4.21.2 SMS/CBS 7-bit Sinhala Alphabet (0x21) extension (SS2) table

				<b>b7</b>	0	0	0	0	1	1	1	1
				<b>b6</b>	0	0	1	1	0	0	1	1
				<b>b5</b>	0	1	0	1	0	1	0	1
<b>b4</b>	<b>b3</b>	<b>b2</b>	<b>b1</b>	<b>1B</b>	0	1	2	3	4	5	6	7
0	0	0	0	0	@ 0040	< 003C	ඇ 0DEA	ංaa 0DF2	। 007C	P 0050	¼ 00BC	△ 26A0
0	0	0	1	1	£ 00A3	= 003D	උ 0DEB	ං♀ 0DF3	A 0041	Q 0051	½ 00BD	∞ 221E
0	0	1	0	2	\$ 0024	> 003E	ඇ 0DEC	ංaa 0D8E	B 0042	R 0052	¾ 00BE	² 00B2
0	0	1	1	3	¥ 00A5	₹ 20B9	ඇ 0DED	ං♀ 0D90	C 0043	S 0053	“ 201C	³ 00B3
0	1	0	0	4	§ 00A7	^ 005E	ඇ 0DEE	ංؒ 0DD0	D 0044	T 0054	” 201D	
0	1	0	1	5	" 0022		ඇ 0DEF	ංؒ 0DD1	E 0045	U 0055	¤ 00A4	
0	1	1	0	6	€ 20AC			ංؒ 0D87	F 0046	V 0056	Ω 03A9	
0	1	1	1	7	% 0025	( 0028	ඇ 0DF4	ංؒ 0D88	G 0047	W 0057	μ 03BC	
1	0	0	0	8	& 0026	) 0029	{ 007B	ංؒ 0DB9	H 0048	X 0058	° 00B0	
1	0	0	1	9	' 0027	 0964	} 007D	ංؒ 0DC6	I 0049	Y 0059	× 00D7	
1	0	1	0	A	FF 000C	 0965	ඇ 0DE7,2044,0DEA	≈ 2248	J 004A	Z 005A	÷ 00F7	
1	0	1	1	B	- 2212	SS3 ----	ඇ 0DE7,2044,0DE8	≠ 2260	K 004B		± 00B1	
1	1	0	0	C	! 0021	ֆ 0DE6	ඇ 0DE9,2044,0DEA	[ 005B	L 004C		% 2030	
1	1	0	1	D	CSI 009B	၏ 0DE7	≤ 2A7D	~ 007E	M 004D		« 00AB	
1	1	1	0	E	- 002D	၏ 0DE8	≥ 2A7E	] 005D	N 004E		» 00BB	
1	1	1	1	F		၏ 0DE9	\ 005C	.	O 004F		,	2019

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.