

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

A.4.21.1 SMS/CBS 7-bit Sinhala Alphabet (0x21) 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	ඌ 0D81	එ 0D93	SP 0020	0 0030	බ 0DB6	ා 0DCF		p 0070
0	0	0	1	1	ො 0D82		ත 0DAD	1 0031	භ 0DB7	ී 0DD2	a 0061	q 0071
0	0	1	0	2	ඃ 0D83	ඹ 0D94	ට 0DA7	2 0032	ම 0DB8	ී 0DD3	b 0062	r 0072
0	0	1	1	3	ආ 0D85	ඹ 0D95	# 0023	3 0033	ය 0DBA	ු 0DD4	c 0063	s 0073
0	1	0	0	4	ආ 0D86	ඹ 0D96	ඵ 0DAE	4 0034	ර 0DBB	ු 0DD6	d 0064	t 0074
0	1	0	1	5	ඉ 0D89	ක 0D9A	ඨ 0DA8	5 0035		ා 0DD8	e 0065	u 0075
0	1	1	0	6	ඊ 0D8A	බ 0D9B	ද 0DAF	6 0036	ල 0DBD	ා 0DDF	f 0066	v 0076
0	1	1	1	7	උ 0D8B	ග 0D9C	ඩ 0DA9	7 0037	ළ 0DC5		g 0067	w 0077
1	0	0	0	8	උ 0D8C	ඝ 0D9D	ධ 0DB0	8 0038		ෙ 0DD9	h 0068	x 0078
1	0	0	1	9	ඝ 0D8D	ඞ 0D9E	ඪ 0DAA	9 0039	ව 0DC0	ේ 0DDA	i 0069	y 0079
1	0	1	0	A	LF 000A	ච 0DA0	* 002A	:	ඪ 0DC3	ෙ 0ddb	j 006A	z 007A
1	0	1	1	B	ඡ 0D8F	SS2 ----	+ 002B	;	ඣ 0DC2		k 006B	ඟ 0D9F
1	1	0	0	C		ඡ 0DA1	, 002C	ඣ 0DAB	ඟ 0DC1	ා 0DDC	l 006C	ඳ 0DA5
1	1	0	1	D	CR 000D	ඡ 0DA2	ත 0DB1	ඣ 0DB4	භ 0DC4	ේ 0DDD	m 006D	ඡ 0DA6
1	1	1	0	E	ඵ 0D91	ක 0DA3	. 002E	ඵ 0DB5		ා 0DDE	n 006E	ඳ 0DAC
1	1	1	1	F	ඵ 0D92	ඳ 0DA4	/ 002F	? 003F		් 0DCA	o 006F	ඳ 0DB3

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

				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	ඒ 0DEA	ආ 0DF2	 007C	P 0050	¼ 00BC	Δ 26A0
0	0	0	1	1	£ 00A3	= 003D	ද 0DEB	ආ 0DF3	A 0041	Q 0051	½ 00BD	∞ 221E
0	0	1	0	2	\$ 0024	> 003E	ඊ 0DEC	ආ 0D8E	B 0042	R 0052	¾ 00BE	² 00B2
0	0	1	1	3	¥ 00A5	₹ 20B9	උ 0DED	ආ 0D90	C 0043	S 0053	“ 201C	³ 00B3
0	1	0	0	4	§ 00A7	^ 005E	ඌ 0DEE	ආ 0DD0	D 0044	T 0054	” 201D	
0	1	0	1	5	" 0022		ඹ 0DEF	ආ 0DD1	E 0045	U 0055	¤ 00A4	
0	1	1	0	6	€ 20AC			ආ 0D87	F 0046	V 0056	Ω 03A9	
0	1	1	1	7	% 0025	(0028	ඹ 0DF4	ආ 0D88	G 0047	W 0057	μ 03BC	
1	0	0	0	8	& 0026) 0029	{ 007B	ඹ 0DB9	H 0048	X 0058	° 00B0	
1	0	0	1	9	' 0027	 0964	} 007D	ඹ 0DC6	I 0049	Y 0059	× 00D7	
1	0	1	0	A	FF 000C	 0965	ඹ 0DE7,2044,0DEA	≈ 2248	J 004A	Z 005A	÷ 00F7	
1	0	1	1	B	- 2212	SS3 ----	ඹ 0DE7,2044,0DE8	≠ 2260	K 004B		± 00B1	
1	1	0	0	C	! 0021	‡ 0DE6	ඹ 0DE9,2044,0DEA	[005B	L 004C		‰ 2030	
1	1	0	1	D	CSI 009B	ඹ 0DE7	≤ 2A7D	~ 007E	M 004D		« 00AB	
1	1	1	0	E	- 002D	ඹ 0DE8	≥ 2A7E] 005D	N 004E		» 00BB	
1	1	1	1	F		ඹ 0DE9	\ 005C	· 00B7	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.