

A.4.1B SMS/CBS 7-bit Tamil Alphabet (0x1B)

A.4.1B.1 SMS/CBS 7-bit Tamil Alphabet (0x1B) 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	ஜ 0B90	SP 0020	0 0030		ஓந் 0BBE	ஓ 0BD0	p 0070		
0	0	0	1	1	ஃ 0B82	05F	த 0BA4	1 0031	ி 0BBF	a 0061	q 0071		
0	0	1	0	2	ஃ 0B83	ஓ 0B92	ஞ 0B9F	2 0032	ங 0BAE	ஓ 0BC0	b 0062	r 0072	
0	0	1	1	3	அ 0B85	ஓ 0B93	# 0023	3 0033	ங 0BAF	ஓந் 0BC1	c 0063	s 0073	
0	1	0	0	4	அ 0B86	ஓள்ள 0B94		4 0034	ஞ 0BB0	ஓந் 0BC2	d 0064	t 0074	
0	1	0	1	5	இ 0B87	க 0B95		5 0035	ங 0BB1		e 0065	u 0075	
0	1	1	0	6	FF 0B88			6 0036	வ 0BB2		f 0066	v 0076	
0	1	1	1	7	உ 0B89			7 0037	ஞ 0BB3		g 0067	w 0077	
1	0	0	0	8	ஊ 0B8A			8 0038	ங 0BB4	ஒ 0BC6	h 0068	x 0078	
1	0	0	1	9		ஏ 0B99		9 0039	வ 0BB5	ஒ 0BC7	i 0069	y 0079	
1	0	1	0	ா	LF 000A	ஃ 0B9A	*	:	ஞ 0BB8	ஒ 0BC8	j 006A	z 007A	
1	0	1	1	ஒ		SS2 ----	+	;	ஏ 0BB7	ஓந் 0BD7	k 006B		
1	1	0	0	ச			,	002C	ஞா 0BA3	ஞ 0BB6	ஒந் 0BCA	l 006C	
1	1	0	1	ஞ	CR 000D	ஜ 0B9C	ங 0BA8	ஞ 0BAA	ஞ 0BB9	ஒந் 0BCB	m 006D		
1	1	1	0	ஏ	ஞ 0B8E		.	002E		ஒந்த 0BCC	n 006E		
1	1	1	1	ஒ	ஞ 0B8F	ஞ 0B9E	/	?		ஒ 0BCD	ஒ 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+FFFFD).

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.4.1B.2). (Note: setting, in the SMS and CBS protocols, base or extension table with reference value 0x10 or larger automatically sets the other tables with the same reference value.) When splitting a message text into submessages, there must be no cut right after an SS2.

0BCD TAMIL SIGN VIRAMA: In Tamil U+0BCD is a combining character with nominal glyph. (However, the closely related Grantha script has a virama that is a conjunct forming control character.)

A.4.1B.2 SMS/CBS 7-bit Tamil Alphabet (0x1B) 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	ஃ 0BEA		 007C
0	0	0	1	1	£ 00A3	= 003D	ஃ 0BEB		A 0041
0	0	1	0	2	\$ 0024	> 003E	ஃ 0BEC		B 0042
0	0	1	1	3	¥ 00A5	₹ 20B9	ஃ 0BED		C 0043
0	1	0	0	4	§ 00A7	^ 005E	ஃ 0BEE		D 0044
0	1	0	1	5	" 0022		ஃ 0BEF		E 0045
0	1	1	0	6	€ 20AC				F 0046
0	1	1	1	7	% 0025	(0028	எ _{ங்கி} 11FFF		G 0047
1	0	0	0	8	& 0026) 0029	{ 007B	ஃ 0BF6	H 0048
1	0	0	1	9	' 0027	 0964	} 007D	எ _{வை} 0BF7	I 0049
1	0	1	0	A	FF 000C	 0965	ஃ/ஃ 0BE7,2044,0BEA	≈ 2248	J 004A
1	0	1	1	B	- 2212	SS3 ----	ஃ/ஃ 0BE7,2044,0BE8	≠ 2260	K 004B
1	1	0	0	C	! 0021	0 0BE6	ஃ/ஃ 0BE9,2044,0BEA	[005B	L 004C
1	1	0	1	D	CSI 009B	ஃ 0BE7	≤ 2A7D	~ 007E	ஃ 004D
1	1	1	0	E	- 002D	ஃ 0BE8	≥ 2A7E] 005D	N 004E
1	1	1	1	F	ஃ 0BA9	ஃ 0BE9	\ 005C	. 00B7	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.