

A.4.15 SMS/CBS 7-bit Gujarati Alphabet (0x15)

A.4.15.1 SMS/CBS 7-bit Gujarati Alphabet (0x15) 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
0	0	0	0	0	ં	એ	સ્પ	૦
0	0	0	1	1	૦૮૧	૦૯૦	૦૦૨૦	૦૦૩૦
0	0	0	1	1	૦૮૨	૦૦૫૯	૦૦૪	૦૦૩૧
0	0	1	0	2	૦૮૩		૦૯૯	૦૦૩૨
0	0	1	1	3	૦૮૫	૦૯૩	#	૩
0	1	0	0	4	૦૮૬	૦૯૪	થ	૪
0	1	0	1	5	૦૮૭	૦૯૫	૬	૫
0	1	1	0	6	૦૮૮	૦૯૬	૪	૬
0	1	1	1	7	૦૮૯	૦૯૭	૮	૭
1	0	0	0	8	૦૮૯	૦૯૮	૯	૮
1	0	0	1	9	૦૮૯૮	૦૯૯	૯	૯
1	0	1	0	ા	૦૦૦૯	૦૯૯	૦૦૨૯	૦૦૩૯
1	0	1	1	બ	૦૮૯૯	૦૦૨૯	*	:
1	1	0	0	ચ	૦૮૯૯	૦૯૯	,	૦૦૨૯
1	1	0	1	દ	૦૦૦૯	૦૯૯	૦૦૨૯	૦૦૩૯
1	1	1	0	એ	૦૮૯૯	૦૯૯	.	૦૦૨૯
1	1	1	1	ફ			/	૦૦૨૯

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+FFFF).

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.4.15.2).

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

0ACD GUJARATI SIGN VIRAMA: This is a control character for forming conjuncts. If the display font does not support the requested conjunct, U+0ACD is a combining character with nominal glyph. In Gujarati, conjuncts are often formed by using the half form of the first consonant, which also may be used final consonants.

A.4.15.2 SMS/CBS 7-bit Gujarati Alphabet (0x15) 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	ય 0AEA	ાર 0AE0	 007C
0	0	0	1	1	£ 00A3	= 003D	૫ 0AEB	૮૯ 0AE1	A 0041
0	0	1	0	2	\$ 0024	> 003E	૬ 0AEC	૭૧ 0AC4	B 0042
0	0	1	1	3	¥ 00A5	₹ 20B9	૭ 0AED	૭૨ 0AE3	C 0043
0	1	0	0	4	§ 00A7	^ 005E	૮ 0AEE		D 0044
0	1	0	1	5	" 0022	૩૦ 0AF1	૮ 0AEF		E 0045
0	1	1	0	6	€ 20AC				F 0046
0	1	1	1	7	% 0025	(0028			G 0047
1	0	0	0	8	& 0026) 0029	{ 007B		H 0048
1	0	0	1	9	' 0027	 0964	} 007D		I 0049
1	0	1	0	A	FF 000C	॥ 0965	૭૪ 0AE7,2044,0AEA	≈ 2248	J 004A
1	0	1	1	B	- 2212	૭૮ ----	૭૮ 0AE7,2044,0AE8	≠ 2260	K 004B
1	1	0	0	C	! 0021	૦ 0AE6	૩૪ 0AE9,2044,0AEA	[005B	L 004C
1	1	0	1	D	CSI 009B	૧ 0AE7	૱૱ 2A7D	~ 007E	M 004D
1	1	1	0	E	- 002D	૨ 0AE8	૱૱ 2A7E] 005D	N 004E
1	1	1	1	F		૩ 0AE9	૱ 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.