Standardization of Sanskrit for Electronic Data Transfer and Screen Representation

Text Encoding Guidelines

During the 8th World Sanskrit Conference, Vienna 1990, a panel was held to discuss the standardization of Sanskrit for electronic data transfer.

8-bit character set for text display

Finally, although the above two provisions cover all essential needs, the panelstill felt that a standard assignment of graphic codes for the display of Sanskrit transliteration would be helpful. An ad hoc committee of interested parties was formed, and two 8-bit "code pages" were designed. One, Classical Sanskrit (CS), for standard use and another, Classical Sanskrit Extended (CSX), which included the former, but also provided for Vedic, MIA, Tamil and some special usages. These code pages take as their point of departure IBM's code page 437, the default set of character codes built into the IBM PC and clones.

These codes were chosen to have minimal impact on the standard IBM PC extended ASCII character set, but they are intended for general use in displaying Indological texts on any machine with an 8-bit (or greater) character set.

Classical Sanskrit (CS)

166 l tilde	ĩ	240 N overdot	Ṅ
167 m overdot	ṁ	241 t underdot	ṭ
224 a macron	ā	242 T underdot	Ṭ
225 not used (normally German *eszett*, ß)		243 d underdot	ḍ
226 A macron	Ā	244 D underdot	Ḍ
227 i macron	ī	245 n underdot	ṇ
228 I macron	Ī	246 N underdot	Ṇ
229 u macron	ū	247 s acute	ś
230 U macron	Ū	248 S acute	Ś
231 r underdot	ṛ	249 s underdot	ṣ
232 R underdot	Ṛ	250 S underdot	Ṣ
233 r underdot macron	ṝ	251 not used (normally the root sign √)	
234 R underdot macron	Ṝ	252 m underdot	ṃ
235 l underdot	ḷ	253 M underdot	Ṃ
236 L underdot	Ḷ	254 h underdot	ḥ
237 l underdot macron	ḹ	255 H underdot	Ḥ
238 L underdot macron	Ḹ		
239 n overdot	ṅ		

Classical Sanskrit Extended (CSX) additions

The following definitions are added to the above Classical Sanskrit character set.

159 r underbar	r̲	199 r underdot grave	ṛ̀
168 a macron breve	ā̆	207 r underdot macron acute	ṝ́
169 i macron breve	ī̆	208 a tilde	ã
170 u macron breve	ū̆	209 i tilde	ĩ
173 n underbar	n̲	210 u tilde	ũ
181 a macron acute	ā́	211 e tilde	ẽ
182 a macron grave	ā̀	212 o tilde	õ
183 i macron acute	ī́	213 e breve	ĕ
184 i macron grave	ī̀	214 o breve	ŏ
189 u macron acute	ū́	215 l underbar	l̲
190 u macron grave	ū̀		
198 r underdot acute	ṛ́		

Key strokes	Pāli Character	Arial Equivalent	Chandra's CSX	Norman	BRM
Alt+ a	ā	à	Alt+ 0224	+0140	+0185
Alt+Shft+ a	Ā	â	Alt+ 0226	+0238	+0176
Alt+ i	ī	ã	Alt+ 0227	+0180	+0186
Alt+Shft+ i	Ī	ä	Alt+ 0228	+0231	+0177
Alt+ u	ū	å	Alt+ 0229	+0232	+0188
Alt+Shft+ u	Ū	æ	Alt+ 0230	+0242	+0178
Alt+ t	ṭ	ñ	Alt+ 0241	+0173	+0169
Alt+Shft+ t	Ṭ	ò	Alt+ 0242	+0233	+0168
Alt+ d	ḍ	ó	Alt+ 0243	+0182	+0140
Alt+Shft+ dd	Ḍ	ô	Alt+ 0244	+0223	+0151
Alt+ l	ḷ	ë	Alt+ 0235	+0194	+0172
Alt+Shft+ l	Ḷ	ì	Alt+ 0236	+0226	+0153
Cont+ m	ṁ	ÿ	Alt+ 0255 (Chandra) Alt+ 0164 (CSX)	+0215	+0174
Cont+Shft+ m	Ṁ	Ñ	Alt+ 0209 (Chandra) Alt+ 0170 (CSX)	+0	+0183
Alt+ m	ṃ	ü	Alt+ 0252	+0181	+0156
Alt+Shft+ m	Ṃ	ý	Alt+ 0253	+0	+0150
Alt+ y	ñ	¤	Alt+ 0164	+0150	+0241
Alt+Shft+ y	Ñ	¥	Alt+ 0165	+0	+0209
Alt+ n	ṇ	õ	Alt+ 0245	+0246	+0133
Alt+Shft+ n	Ṇ	ö	Alt+ 0246	+0	+0164
Cont+ n	ṅ	ï	Alt+ 0239	+0186	+0137
Cont+Shft+ n	Ṅ	ð	Alt+ 0240	+0	+0162

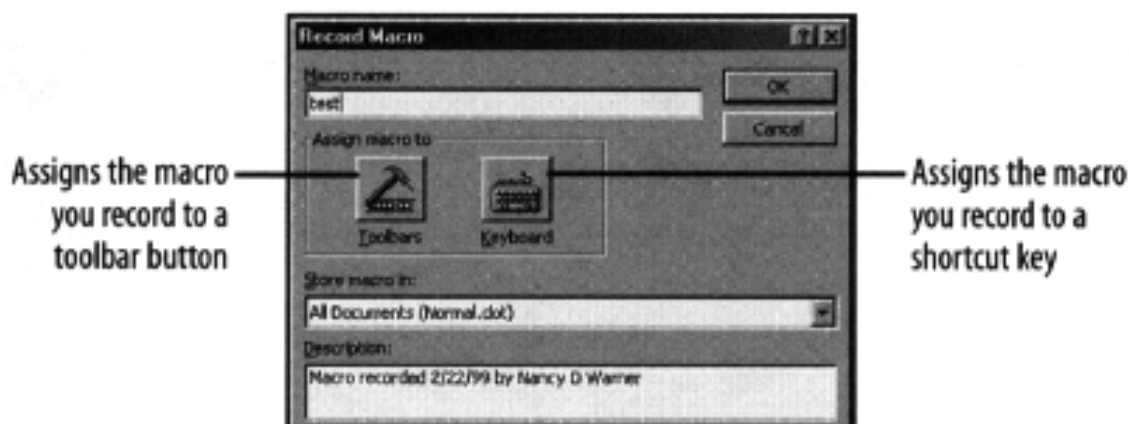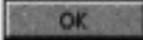| +0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11	+12	+13	+14	+15
0 •	1 •	2 •	3 •	4 •	5 •	6 •	7 •	8 •	9 •	10 •	11 •	12 •	13 •	14 •	15 •
16 •	17 •	18 •	19 •	20 •	21 •	22 •	23 •	24 •	25 •	26 •	27 •	28 •	29 •	30 •	31 •
32	33 !	34 "	35 #	36 $	37 %	38 &	39 '	40 (	41)	42 *	43 +	44 ,	45 -	46 .	47 /
48 0	49 1	50 2	51 3	52 4	53 5	54 6	55 7	56 8	57 9	58 :	59 ;	60 <	61 =	62 >	63 ?
64 @	65 A	66 B	67 C	68 D	69 E	70 F	71 G	72 H	73 I	74 J	75 K	76 L	77 M	78 N	79 O
80 P	81 Q	82 R	83 S	84 T	85 U	86 V	87 W	88 X	89 Y	90 Z	91 [	92 \	93]	94 ^	95 _
96 `	97 a	98 b	99 c	100 d	101 e	102 f	103 g	104 h	105 i	106 j	107 k	108 l	109 m	110 n	111 o
112 p	113 q	114 r	115 s	116 t	117 u	118 v	119 w	120 x	121 y	122 z	123 {	124 |	125 }	126 ~	127 •
128 •	129 '	130 ‚	131 ƒ	132 „	133 …	134 †	135 ‡	136 ˆ	137 ‰	138 Š	139 ‹	140 Œ	141 •	142 Ž	143 •
144 •	145 '	146 '	147 "	148 "	149 •	150 –	151 —	152 ˜	153 ™	154 š	155 ›	156 œ	157 •	158 ž	159 Ÿ
160 •	161 ị	162 ¢	163 £	164 ñ	165 Ñ	166 ¦	167 §	168 ¨	169 ©	170 ª	171 «	172 ¬	173	174 ®	175 ¯
176 °	177 ±	178 ²	179 ³	180 ´	181 µ	182 ¶	183 ·	184 ¸	185 ¹	186 º	187 »	188 ¼	189 ½	190 ¾	191 ¿
192 À	193 Á	194 Â	195 Ã	196 Ä	197 Å	198 Æ	199 Ç	200 È	201 É	202 Ê	203 Ë	204 Ì	205 Í	206 Î	207 Ï
208 Ð	209 Ṁ	210 Ò	211 Ó	212 Ô	213 Õ	214 Ö	215 ×	216 Ø	217 Ù	218 Ú	219 Û	220 Ü	221 Ý	222 Þ	223 ß
224 ā	225 á	226 Ā	227 ī	228 Ī	229 ū	230 Ū	231 ç	232 è	233 é	234 ê	235 ị	236 Ḷ	237 ī	238 î	239 ṅ
240 Ṅ	241 ṭ	242 Ṭ	243 ḍ	244 Ḍ	245 ṇ	246 Ṇ	247 ÷	248 ø	249 ù	250 û	251 ū	252 ṃ	253 Ṃ	254 ṃ	255 •

Roman Pali / Sanskrit Encodings

Values are in hexidecimal / decimal

	Unicode Name	Unicode	VRI	CSX	Norman
Ñ	LATIN CAPITAL LETTER N WITH TILDE	00D1 / 209	D1 / 209	A5 / 165	84 / 132
ñ	LATIN SMALL LETTER N WITH TILDE	00F1 / 241	F1 / 241	A4 / 164	96 / 150
Ā	LATIN CAPITAL LETTER A WITH MACRON	0100 / 256	BE / 190	E2 / 226	EE / 238
ā	LATIN SMALL LETTER A WITH MACRON	0101 / 257	B1 / 177	E0 / 224	8C / 140
Ī	LATIN CAPITAL LETTER I WITH MACRON	012A / 298	BF / 191	E4 / 228	E7 / 231
ī	LATIN SMALL LETTER I WITH MACRON	012B / 299	B2 / 178	E3 / 227	B4 / 180
Ś	LATIN CAPITAL LETTER S WITH ACUTE	015A / 346	E6 / 230	F8 / 248	EA / 234
ś	LATIN SMALL LETTER S WITH ACUTE	015B / 347	9C / 156	F7 / 247	A7 / 167
Ū	LATIN CAPITAL LETTER U WITH MACRON	016A / 362	D0 / 208	E6 / 230	F2 / 242
ū	LATIN SMALL LETTER U WITH MACRON	016B / 363	B3 / 179	E5 / 229	E8 / 232
Ḍ	LATIN CAPITAL LETTER D WITH DOT BELOW	1E0C / 7692	DE / 222	F4 / 244	DF / 223
ḍ	LATIN SMALL LETTER D WITH DOT BELOW	1E0D / 7693	B9 / 185	F3 / 243	B6 / 182
Ḥ	LATIN CAPITAL LETTER H WITH DOT BELOW	1E24 / 7716	89 / 137	FF / 255	DE / 222
ḥ	LATIN SMALL LETTER H WITH DOT BELOW	1E25 / 7717	83 / 131	FE / 254	FA / 250
Ḷ	LATIN CAPITAL LETTER L WITH DOT BELOW	1E36 / 7734	FC / 253	EC / 236	E2 / 226
ḷ	LATIN SMALL LETTER L WITH DOT BELOW	1E37 / 7735	BC / 188	EB / 235	C2 / 194
Ṃ	LATIN CAPITAL LETTER M WITH DOT BELOW	1E42 / 7746	FE / 254	FD / 253	C8 / 200
ṃ	LATIN SMALL LETTER M WITH DOT BELOW	1E43 / 7747	BD / 189	FC / 252	B5 / 181
Ṅ	LATIN CAPITAL LETTER N WITH DOT ABOVE	1E44 / 7748	A9 / 169	F0 / 240	DC / 220
ṅ	LATIN SMALL LETTER N WITH DOT ABOVE	1E45 / 7749	AA / 170	EF / 239	BA / 186
Ṇ	LATIN CAPITAL LETTER N WITH DOT BELOW	1E46 / 7750	F0 / 240	F6 / 246	EF / 239
ṇ	LATIN SMALL LETTER N WITH DOT BELOW	1E47 / 7751	BA / 186	F5 / 245	F6 / 246
Ṛ	LATIN CAPITAL LETTER R WITH DOT BELOW	1E5A / 7770	AE / 174	E8 / 232	---
ṛ	LATIN SMALL LETTER R WITH DOT BELOW	1E5B / 7771	A5 / 165	E7 / 231	A8 / 168
Ṣ	LATIN CAPITAL LETTER S WITH DOT BELOW	1E62 / 7778	C6 / 198	F0 / 250	DD / 221
ṣ	LATIN SMALL LETTER S WITH DOT BELOW	1E63 / 7779	8C / 140	F9 / 249	B7 / 183
Ṭ	LATIN CAPITAL LETTER T WITH DOT BELOW	1E6C / 7788	DD / 221	F2 / 242	E9 / 233
ṭ	LATIN SMALL LETTER T WITH DOT BELOW	1E6D / 7789	B5 / 181	F1 / 241	AD / 173

All of the characters above are located in the Unicode ranges C1 Controls and Latin-1 Supplement, (Ñ and ñ), Latin Extended A, and Latin Extended Additional.

Create a Macro

1. Choose **Tools, Macro, Record New Macro** to open the Record Macro dialog box.

2. Type a name in the **Macro name** text box.

Assigns the macro you record to a toolbar button ————

———— Assigns the macro you record to a shortcut key

3. Click the **OK** [OK] button. The Macro toolbar appears with the **Stop Recording** ■ and **Pause Recording** button.

4. Perform any tasks that you want the macro to record and click the **Stop Recording** ■ button when finished.

TIP

Macros are also items that you can add to your toolbars to make it easy to launch them. Read the Toolbars section and follow along the steps of adding a button to your toolbar, but choose **Macros** from the **Categories** list. Then choose the specific macro in the **Commands** list.

Run a Macro

1. Press the [Alt]+[F8] keys to open the Macros dialog box.

2. Double-click the **Macro name** and the macro will run.

Modify or Delete a Macro

1. Choose **Tools, Macro, Macros** to open the Macro dialog box.

2. Click the **Macro name** and click the appropriate button to perform the modification you desire:

   ■ **Edit**—Opens the Visual Basic editor to the macro code. Make any changes and click the **Close** ☒ button to return to your document.