

UTR#50 Vertical_Orientation

Assignments for Unicode 9.0 Characters

To: UTC
 From: Koji Ishii (Google Inc.), Laurențiu Iancu (Microsoft), Eric Muller (Amazon.com, Inc.)
 Date: 2016-07-28

Overview

This document explains assignments of Vertical_Orientation for new characters in Unicode 9.0 in the [proposed update of UTR#50 revision 16](#).

Changes from current defaults are marked with reserved values in parenthesis in this table, and they documented in the Modifications section of UTR#50.

Data

Code points	Value (reserved)	Name and Notes
08B6..08BD 08D4..08E1 08E2 0C80 0D4F 0D54..0D56 0D58..0D5E 0D76..0D78	R	ARABIC LETTER BEH WITH SMALL MEEM ABOVE..ARABIC LETTER AFRICAN NOON ARABIC SMALL HIGH WORD AR-RUB..ARABIC SMALL HIGH SIGN SAFHA ARABIC DISPUTED END OF AYAH KANNADA SIGN SPACING CANDRABINDU MALAYALAM SIGN PARA MALAYALAM LETTER CHILLU M..MALAYALAM LETTER CHILLU LLL MALAYALAM FRACTION ONE ONE-HUNDRED-AND-SIXTIETH..MALAYALAM FRACTION ONE FIFTH MALAYALAM FRACTION ONE SIXTEENTH..MALAYALAM FRACTION THREE SIXTEENTHS Consistent with the vo value for existing characters in those scripts and blocks.
1C80..1C88	R	CYRILLIC SMALL LETTER ROUNDED VE..CYRILLIC SMALL LETTER UNBLENDED UK Proposed assignment is for the whole new block Cyrillic

		Extended-C (U+1C80..U+1C8F). Consistent with existing Cyrillic characters in other blocks.
1DFB	R	COMBINING DELETION MARK Consistent with existing diacritics in the same block and with Newa, which uses it.
23FB..23FE	U	POWER SYMBOL..POWER SLEEP SYMBOL Consistent with adjacent UI symbols in Miscellaneous Technical.
2E43..2E44 A7AE A8C5 1018D..1018E	R	DASH WITH LEFT UPTURN..DOUBLE SUSPENSION MARK LATIN CAPITAL LETTER SMALL CAPITAL I SAURASHTRA SIGN CANDRABINDU GREEK INDICATION SIGN..NOMISMA SIGN Consistent with similar characters.
104B0..104D3 104D8..104FB	R	OSAGE CAPITAL LETTER A..OSAGE CAPITAL LETTER ZHA OSAGE SMALL LETTER A..OSAGE SMALL LETTER ZHA Assignment inspired by Cherokee and proposed for the whole new block Osage (U+104B0..U+104FF).
1123E	R	KHOJKI SIGN SUKUN Consistent with existing characters.
11400..11459 1145B 1145D	R	NEWA LETTER A..NEWA DIGIT NINE NEWA PLACEHOLDER MARK NEWA INSERTION SIGN Assignment inspired by Devanagari and proposed for the whole new block Newa (U+11400..U+1147F).
11660..1166C	R	MONGOLIAN BIRGA WITH ORNAMENT..MONGOLIAN TURNED SWIRL BIRGA WITH DOUBLE ORNAMENT Consistent with existing birgas and with Mongolian as a whole, and following the more recent convention with code-chart glyphs shown as for horizontal layout. Proposed assignment is for the whole new block Mongolian Supplement (U+11660..U+1167F).
11C00..11C08	R	BHAIKSUKI LETTER A..BHAIKSUKI LETTER VOCALIC L

<p>11C0A..11C36 11C38..11C45 11C50..11C6C</p>		<p>BHAIKSUKI LETTER E..BHAIKSUKI VOWEL SIGN VOCALIC L BHAIKSUKI VOWEL SIGN E..BHAIKSUKI GAP FILLER-2 BHAIKSUKI DIGIT ZERO..BHAIKSUKI HUNDREDS UNIT MARK</p> <p>Assignment inspired by Devanagari and Sharada, and proposed for the whole new block Bhaiksuki (U+11C00..U+11C6F).</p>
<p>11C70..11C8F 11C92..11CA7 11CA9..11CB6</p>	R	<p>MARCHEN HEAD MARK..MARCHEN LETTER A MARCHEN SUBJOINED LETTER KA..MARCHEN SUBJOINED LETTER ZA MARCHEN SUBJOINED LETTER YA..MARCHEN SIGN CANDRABINDU</p> <p>Assignment inspired by Tibetan, with similar encoding model with subjoined consonants for forming stacked conjuncts. Assignment proposed for the whole new block Marchen (U+11C70..U+11CBF).</p>
<p>16FE0</p>	U (R)	<p>TANGUT ITERATION MARK</p> <p>Assignment proposed for consistency with the two Tangut blocks (see below). The question is whether a default of Upright should be defined now for the whole new block Ideographic Symbols and Punctuation (U+16FE0..U+16FFF) -or- the default should stay Rotated and vo values be defined piecemeal as new characters are encoded. The next future character in that block will be a Nushu iteration mark (U+16FE1; see the pipeline), which will probably be Upright. Comparing with the CJK Symbols and Punctuation block, most characters there are Upright, with brackets and dashes being Tr, and a couple Tu. From that perspective, it might be less churn to go ahead and define the default to be Upright for the whole block.</p>
<p>17000..187EC 18800..18AF2</p>	U (R)	<p>TANGUT IDEOGRAPH-17000..TANGUT IDEOGRAPH-187EC TANGUT COMPONENT-001..TANGUT COMPONENT-755</p> <p>Assignment based on CJK, as Tangut is a siniform ideographic script, and proposed for the whole new blocks Tangut (U+17000..U+187FF) and Tangut Components (U+18800..U+18AFF).</p>
<p>1E000..1E006 1E008..1E018</p>	R	<p>COMBINING GLAGOLITIC LETTER AZU..COMBINING GLAGOLITIC LETTER ZHIVETE</p>

<p>1E01B..1E021 1E023..1E024 1E026..1E02A</p>		<p>COMBINING GLAGOLITIC LETTER ZEMLJA..COMBINING GLAGOLITIC LETTER HERU COMBINING GLAGOLITIC LETTER SHTA..COMBINING GLAGOLITIC LETTER YATI COMBINING GLAGOLITIC LETTER YU..COMBINING GLAGOLITIC LETTER SMALL YUS COMBINING GLAGOLITIC LETTER YO..COMBINING GLAGOLITIC LETTER FITA</p> <p>Consistent with existing Glagolitic and proposed for the whole new block Glagolitic Supplement (U+1E000..U+1E02F).</p>
<p>1E900..1E94A 1E950..1E959 1E95E..1E95F</p>	R	<p>ADLAM CAPITAL LETTER ALIF..ADLAM NUKTA ADLAM DIGIT ZERO..ADLAM DIGIT NINE ADLAM INITIAL EXCLAMATION MARK..ADLAM INITIAL QUESTION MARK</p> <p>Assignment inspired by N'Ko and proposed for the whole new block Adlam (U+1E900..U+1E95F).</p>
<p>1F19B..1F1AC</p>	U	<p>SQUARED THREE D..SQUARED VOD</p> <p>Consistent with existing enclosed alphanumerics in the same block.</p>
<p>1F23B</p>	U	<p>SQUARED CJK UNIFIED IDEOGRAPH-914D</p> <p>Consistent with existing enclosed ideographs in the same block.</p>
<p>1F57A 1F5A4 1F6D1..1F6D2 1F6F4..1F6F6 1F919..1F91E 1F920..1F927 1F930 1F933..1F93E 1F940..1F94B 1F950..1F95E 1F985..1F991</p>	U	<p>MAN DANCING BLACK HEART OCTAGONAL SIGN..SHOPPING TROLLEY SCOOTER..CANOE CALL ME HAND..HAND WITH INDEX AND MIDDLE FINGERS CROSSED FACE WITH COWBOY HAT..SNEEZING FACE PREGNANT WOMAN SELFIE..HANDBALL WILTED FLOWER..MARTIAL ARTS UNIFORM CROISSANT..PANCAKES EAGLE..SQUID</p> <p>Consistent with existing emoji in the blocks in that area.</p>