L2/17-038

Indic positional category for Javanese cakra

Norbert Lindenberg, 2017-01-21

This document provides samples for the bug report "Incorrect Indic positional category for Javanese consonant sign cakra", which I submitted to the Unicode Consortium on 2016-08-10. The samples show cakra in a book printed in 1903 as well as rendered by several digital fonts.

Bug report

The Unicode 9.0 data file IndicPositionalCategory.txt [1] gives the character JAVANESE CONSONANT SIGN CAKRA the positional category Right.

In reality, the cakra is rendered either below its base consonant (in different shapes, many of which wrap around the left side of the base consonant) or, less commonly, isolated to the left of its base consonant. In some styles, a cakra glyph may initially start out to the right from the bottom right corner of its base consonant, but then it always turns down and to the left.

Using the positional category Right also causes problems with the OpenType Universal Shaping Engine [2]. There's another Javanese medial consonant with the same positional category, JAVANESE CONSONANT SIGN PENGKAL, and the USE allows only one medial consonant of each positional category per cluster. As Javanese expert Aditya Bayu Perdana informed me, the "cakra+pengkal combination is actually fairly common in Sanskrit and Kawi literature and are well attested." Examples he provided from a Bharatayuddha epic printed in 1903 are available on request. The current combination of Unicode data and USE specification does not allow the cakra+pengkal combination.

The positional category for JAVANESE CONSONANT SIGN CAKRA should therefore be changed to Bottom, both to better represent the actual positioning of cakra and to enable the cakra +pengkal combination in the USE.

[1] http://www.unicode.org/Public/9.0.0/ucd/IndicPositionalCategory.txt
[2] http://www.microsoft.com/typography/OpenTypeDev/USE/intro.htm#clustervalidation

Sources and fonts used in the samples

Bharatayuddha samples are from pages ꧓꧔, ꧗꧒, and ꧑꧗ of J. G. H. Gunning: Bhârata-yuddha: Oudjavaansch Heldendicht. Martinus Nijhoff 1903. https://upload.wikimedia.org/wikipedia/commons/6/60/Bharata-Yuddha_oudjavaansch_heldendicht.pdf

Tuladha Jejeg 2.01 by R.S. Wihananto is based on Graphite and shown as rendered by Firefox 50.1.0 on OS X 10.11.6.

Yogya 1.0, developed by Lindenberg Software using glyphs by Aditya Bayu Perdana, is based on Apple Advanced Typography and shown as rendered by Safari 10.0.2 on Mac OS 10.11.6.

Javanese Text 1.07, developed by Tiro Typeworks for Microsoft, is based on OpenType with support for Javanese shaping and shown as rendered by Microsoft Edge 38.14393.0.0 on Windows 10 version 1607.

Noto Sans Javanese 1.06, developed by Monotype for Google, is based on OpenType with support for Javanese shaping and shown as rendered by Google Chrome 55.0.2883.95 on OS X 10.11.6.

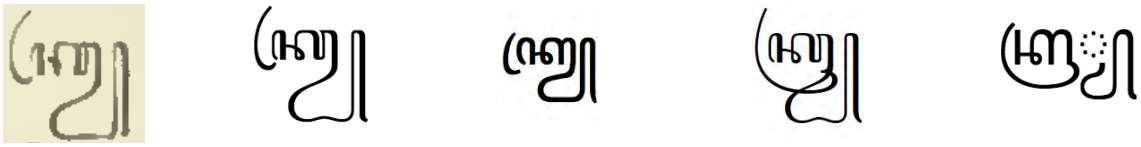
Samples

Bharatayuddha	Tuladha Jejeg Firefox	Yogya Safari	Javanese Text Edge	Noto Sans Javan. Chrome
				

Dotted circle with cakra. This doesn't occur in normal documents. The fonts use cakra shapes that they'd also use with consonants that don't have a leg to attach cakra to. Edge adds a dotted circle for unknown reasons.

Javanese letter ka with cakra. Cakra starts below or attaches to the right leg of the base consonant and wraps around the left side. Adjusted for the different widths of base consonants, this is the most common shape for cakra.

Dotted circle with pengkal, for reference.



Javanese letter ka with cakra and pengkal. Normal behavior is for cakra to take on its isolated form to the left of the base consonant in order to make room for pengkal. Edge omits the dotted circle that the current USE specification and Unicode data require.

Javanese letter sa with conjunct ta (pangkon + ta), cakra, pengkal, tarung, and layar. This syllable is more complex than most, and only one of the tested fonts matches the print sample. Edge omits the dotted circle that the current USE specification and Unicode data require.
