

**Universal Multiple-Octet Coded Character Set
International Organization for Standardization**

Doc Type: UTC

Title: Proposal to allow negative residual strokes in kRS* (radical/stroke) fields of Unihan

Source: Ken Lunde (Adobe)

Status: Individual Contribution

Action: Discussion by UTC

Date: 2015-11-04

A small—or at least limited—number of characters in the [Unihan Database](#) have representative glyphs that are the same as—or very similar to—the indexing radical to which they are assigned, but have one or more strokes obviously missing, which results in a reduction of their total number of strokes. The convention has been to use 0 (zero) as the number of residual strokes for such characters. I propose that negative residual stroke values be allowed, and that the Unihan Database entries for existing characters be adjusted to reflect this, at least for obvious cases, by modifying existing values or adding new ones.

Examples for demonstrating the benefits of this proposal are the following Extension B and E characters: U+225A9 (B), U+29C0A (B), U+29C0B (B), and U+2C09B (E). The relevant Unihan Database entries are shown below:

U+225A9	kRSUnicode	61.0
U+225A9	kRSKangXi	61.0
U+29C0A	kRSUnicode	191.0
U+29C0A	kRSKangXi	191.0
U+29C0B	kRSUnicode	191.0
U+29C0B	kRSKangXi	191.0
U+2C09B	kRSUnicode	75.0

If this proposal is accepted, the kRSUnicode and kRSKangXi values for these characters would be adjusted to become to the following:

U+225A9	kRSUnicode	61.-1
U+225A9	kRSKangXi	61.-1
U+29C0A	kRSUnicode	191.-5
U+29C0A	kRSKangXi	191.-5
U+29C0B	kRSUnicode	191.-5
U+29C0B	kRSKangXi	191.-5
U+2C09B	kRSUnicode	75.-2

Below are code chart excerpts for these four characters so that the obviously absent strokes can be seen:

225A9

心 61.0

UCS2003

T5-2131

29C0A

𠄎 191.0

UCS2003

GHZ-10014.02

T6-233D

29C0B

𠄎 191.0

UCS2003

GHZ-10015.02

T6-233C

2C09B

木 75.0

JK-58711

The greatest benefit of allowing negative residual strokes is that doing so explicitly acknowledges that strokes which would normally constitute the indexing radical are intentionally missing, and not a representative glyph error.

If this proposal is accepted, the "Syntax" and "Description" fields of the following six UAX #38 properties would require adjustment to allow negative residual stroke values: *kRSAdobe_Japan1_6*, *kRSJapanese*, *kRSKangXi*, *kRSKanWa*, *kRSKorean*, and *kRSUnicode*.

BTW, the *kTotalStrokes* values for U+29C0A and U+29C0B, both set to 10, are incorrect, and both should instead be set to 5. ☺

That is all.