

Universal Multiple-Octet Coded Character Set
International Organization for Standardization

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG

Title: Allowing negative residual strokes in all six kRS* properties of the Unihan Database

Source: Unicode Consortium

Status: Liaison Contribution

Action: For consideration by IRG

Date: 2015-11-04

A small—or at least limited—number of CJK Unified Ideographs 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 entire 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. Negative residual stroke values are now allowed, and the [Unihan Database](#) entries for existing characters will be adjusted to reflect this, at least for obvious cases, by modifying existing values.

Examples for demonstrating the benefits of this change 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

The kRSUnicode and kRSKangXi values for these characters will 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

𠂔 191.0

UCS2003

GHZ-10014.02

T6-233D

29C0B

𠂔 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.

As a result of this change, the "Syntax" and "Description" fields of the following six [UAX #38](#) properties will be adjusted to allow negative residual stroke values: *kRSAdobe_Japan1_6*, *kRSJapanese*, *kRSKangXi*, *kRSKanWa*, *kRSKorean*, and *kRSUnicode*.

When sorting, negative residual stroke values should be replaced with 0 (zero) to prevent characters with such values from sorting before the characters that represent the radical itself. This change will also be reflected in UAX #38.

That is all.