

# Universal Coded Character Set UCS

**Title:** Proposal to allow only one `kTotalStrokes` property value

**Source:** Ken Lunde, IRG Convenor

**Distribution:** IRG Member Bodies & Experts

**Date:** 2024-08-03

This document proposes that the informative `kTotalStrokes` Unihan database property be changed to allow only a single property value, specifically the value that adheres to the IRG’s method of counting strokes. All other total stroke counts should be reflected in the provisional `kAlternateTotalStrokes` property that was added in Unicode Version 15.0 (2022) for this specific purpose.

## Background

The `kTotalStrokes` property was added to the Unihan database in Unicode Version 3.1 (2001) with the following description:

The total number of strokes in the character (including the radical)

The description changed to the following in Unicode Version 4.1 (2005):

The total number of strokes in the character (including the radical). This value is for the character as drawn in the Unicode charts.

Although **space** was its delimiter from the beginning, meaning that allowing multiple property values was implicit, it wasn’t until Unicode Version 6.1 (2012) that its description explicitly mentioned multiple property values:

The total number of strokes in the character (including the radical), that is, the stroke count most commonly associated with the character in modern text using customary fonts.

*Multiple Value Order:* When there are two values, then the first is preferred for zh-Hans (CN) and the second is preferred for zh-Hant (TW). When there is only one value, it is appropriate for both.

## Proposal

The number of residual strokes that are specified for the informative `kRSUnicode` property follows the IRG’s method of counting strokes, at least for the most recent IRG working sets, so applying the same convention to the informative `kTotalStrokes` property makes sense. All other total stroke counts should be reflected using the provisional `kAlternateTotalStrokes` property, which is its purpose.

## Discussion Points

The very first multiple `kTotalStrokes` property value was added in Unicode Version 6.3 (**2013**), specifically for U+8303 范 with 8 and 9 as the property values, which remained the only ideograph with multiple `kTotalStrokes` property values through Unicode Version 13.0 (2020). I once asked John Jenkins (RIP) about this single ideograph with two `kTotalStrokes` property values, and he told me that its was added merely for validation purposes.

A second `kTotalStrokes` property value was added to U+2A060 鴻 in Unicode Version 14.0 (**2021**), but because the second property value, 17, was intended for Vietnamese, it should have been moved to the `kAlternateTotalStrokes` property in Unicode Version 15.0 (2022) as follows:

17:V

A second `kTotalStrokes` property value was added to U+9AA8 骨 in Unicode Version 15.0 (**2022**), specifically 10. This was intended to be synchronized with the following corresponding entry in the

*kAlternateTotalStrokes* property that was added in the same version of the Unicode Standard:

10:HJKPV

A second *kTotalStrokes* property value is being added to U+2010E 𠄞 in Unicode Version 16.0 (2024), specifically 1. This is intended to be synchronized with U+31E1 𠄞 CJK STROKE HZZZG in the [CJK Strokes](#) block. All characters in that particular block are considered to have only one stroke.

In summary, among the nearly 100,000 CJK Unified Ideographs in the Unicode Standard, only the following four have more than one *kTotalStrokes* property value, one of which should be moved to the *kAlternateTotalStrokes* property (highlighted in red):

U+8303	kTotalStrokes	8 9
U+9AA8	kTotalStrokes	9 10
U+2010E	kTotalStrokes	1 2
U+2A060	kTotalStrokes	18 <b>17</b>

In other words, there has been almost no attempt to add a second property value to the *kTotalStrokes* property.

## Next Steps

If the IRG agrees with this proposed change to the *kTotalStrokes* property, we can explore ways to retroactively adjusting the existing *kTotalStrokes* property values to adhere to the IRG's conventions. The syntax of the *kAlternateTotalStrokes* property will also need to be adjusted to accommodate the "G" and "T" source identifiers. I will also forward this proposal to the UTC for consideration.

That is all.