



344

New Unicode Character Property Equivalent_Unified_Ideograph

Closing Date:
2017.05.01

Status: Open

Originator: UTC

Informal Discussion: Unicode Mail List ([Join](#))

Formal Feedback: [Contact Form](#)

Description of Issue:

A new character property *Equivalent_Unified_Ideograph* is proposed for addition to Unicode 10.0. This property associates (where possible) the 365 characters in the CJK Radicals Supplement, Kangxi Radicals, and CJK Strokes blocks to an appropriate CJK Unified Ideograph.

An Ideographic Description Sequence (IDS) is used to analyze CJK Unified Ideographs by breaking them down into their structure and components. An IDS is defined in terms of “Ideographic | Radical | CJK_Stroke” (*The Unicode Standard*, Section 18.2). However, there is a certain lack of data about the Radicals and CJK_Strokes. There is a mapping from certain of the radicals to strokes and Unified Ideographs in [CJKRadicals.txt](#), but it is incomplete, and also sometimes inconsistent with the usage in the [UCA](#) and [NamesList.txt](#). In addition to IDSes, there are also other use cases which would benefit from having more complete data.

The UTC is proposing to also add corresponding derived property values for kRSUnicode and kTotalStrokes (see [UAX #38: Unicode Han Database \(Unihan\)](#)) based on the Equivalent_Unified_Ideograph values.

The proposed mappings for Equivalent_Unified_Ideograph can be reviewed in the [proposed property data](#), and the proposed additional values for kRSUnicode and kTotalStrokes can be reviewed in the [proposed derived property data](#).

Draft data updated 2017-01-30.

How to Provide Feedback: For information about how to discuss this Public Review Issue and how to supply formal feedback, please see the [feedback and discussion instructions](#). The accumulated feedback received so far on this issue is shown below, or you can look at a [full page view](#). Feedback is reviewed by the relevant committee according to their meeting schedule.