

Title: Proposal to add two new provisional Unihan database properties

Author: Ken Lunde

Date: 2026-03-05

This document is a proposal to add the following two new provisional properties to the Unihan database in Unicode Version 18.0: **kJapaneseNewVariant** and **kJapaneseOldVariant**.

These properties serve as the Japanese counterparts of the provisional *kSimplifiedVariant* and *kTraditionalVariant* properties that were added to the Unihan database in Unicode Version 2.0 (1996), which are intended only for describing simplified/traditional relationships as they apply to Chinese.

Japanese also uses simplified forms, and maintains simplified/traditional relationships. Japan refers to the simplified and traditional forms as 新字体 (*shinjitai*) and 旧字体 (*kyūjitai*), respectively, which translate as *new forms* and *old forms*, respectively. This is the fundamental reason as to why the proposed property names use **New** and **Old**, not **Simplified** and **Traditional**.

The table below demonstrates the current shortcoming in the Unihan database by showing the simplified forms of U+528D 劍, and how the simplified/traditional relationships are reflected in the Unihan database only for Chinese:

Traditional Form	Simplified Form		Unihan Database
	Chinese	Japanese	
U+528D 劍	U+5251 剑	U+5263 剣	U+528D kSimplifiedVariant U+5251 U+5251 kTraditionalVariant U+528D

The following tables provide the fields for these properties as they would appear in UAX #38:

Field	Text
Property	kJapaneseNewVariant
Status	Provisional
Category	Variants
Introduced	18.0
Delimiter	N/A
Syntax	U\[23]?[0-9A-F]{4}
Description	The Unicode value for the new Japanese form (新字体) of this ideograph.

Field	Text
Property	kJapaneseOldVariant
Status	Provisional
Category	Variants
Introduced	18.0
Delimiter	space
Syntax	U\[23]?[0-9A-F]{4}
Description	One or more Unicode values for the old Japanese forms (旧字体) of this ideograph.

Property data

The proposed property data is based on the official simplified/traditional mappings as shown in the *Jōyō Kanji* (常用漢字) and *Jinmei-yō Kanji* (人名用漢字) tables, along with 28 simplified/traditional pairs that were derived from the relationship between the JIS X 0208 and JIS X 0212 standards.

The proposed property data uses single property values for the **kJapaneseNewVariant** property, and multiple property values for the **kJapaneseOldVariant** property. This is the reason why the *Delimiter* is **N/A** and **space**, respectively.

Data files

The *kJapaneseNewVariant.txt* and *kJapaneseOldVariant.txt* data files, which are PDF attachments, provide everything that is necessary for adding these two provisional properties and their property data to the Unihan Database, and for adding a full description of the properties to [UAX #38](#), *Unicode Han Database (Unihan)*. Both files include at the end human-readable mappings between the two forms.

That is all.