Doc Type: ISO/IEC JTC1/SC2/WG2/IRG
Title: IRG N2776 Update of V-Source Representative Glyphs
Source: Vietnam
Status: National Body Proposal
Authors: Eiso Chan, Lee Collins, Ngô Trung Việt
Action: For consideration by the IRG
Date: 2024-12-30

Reviewers of the latest published Unicode charts have found minor problems with a small number of V-Source characters. The problems are minor, but worth fixing since they will clarify the structure and meaning of the characters involved.

Detailed Rationale

1. V2-7A32 / U+23AF2

As shown above, the reference glyph for this character currently has 母 (U+6BCD, "mother") as its top, semantic component. The phonetic element is 仃 (Sino-Vietnamese: *đinh*) to represent the native Vietnamese reading: *đừng*, meaning "do not!".

The source originally cited for V2-7A32 is "Bảng Tra Chữ Nôm", p. 49, shown below, where we see that the top element is a hand written form that can represent 母 or 毋.

Semantically, we would expect the top element in U+23AF2 to be 毋 (U+6BCB. "do not!"). We have found no other cases of 母 being used with this semantic. It is possible that the current shape is due to the choice of the wrong semantic element when the glyph was redesigned for a printed Kai-style font.

1

The semantically correct form is attested on p. 417 of the dictionary *Tự Điên Chứ Nôm Trích Dẫn* shown below:

We propose that the new design follow that shown above, since this matches the stand-alone form of 毋 in the Vietnamese reference font.

Although this change will reduce the strokes in the glyph by one, there will be no change in the Unihan data. This is because the values for kRSUnicode and kTotal strokes are already based on the stroke counts for U+6BCB. This is also further evidence that the intended shape was U+6BCB.

2. V3-3570 / U+291AC

霻 (U+291AC) is read "mống" and means "rainbow". We would expect the phonetic element to be 蒙 (U+2B387) with the Sino-Vietnamese reading "mông". U+2B387 is the phonetic in a number of characters with similar readings:

U+2C7C3	蠓	mùng
U+232E1	曚	mông
U+2F8AF	懞	mong
… etc.		

As far as we know, the current design of U+291AC, with 田业豕 on bottom, is unique in Vietnamese usage. We can find no justification for this shape. We propose to remove the superfluous stroke 一 and join the the top four strokes to 豕. This is a minor change that will help clarify the structure of U+291AC. The removal of one stroke will require a change in the Unihan data.

3. V2-7F33 / U+2582D

2582D
禾 115.9
UCS2003 V2-7F33

The right side component of this character, 禾 (U+79BE), is incorrectly designed using the shape of 禾 that would normally appear on the left side. We propose modifying the 禾 to have a full final stroke.

4. Normalization adjustments.

We have identified a small number of characters that do not conform to the rules in *IRGN2673 Vietnamese Hán-Nôm Normalization Guidelines*. We propose to bring the following characters into conformance.

23597
木 75.9
UCS2003 V2-7938

23DEF
水 85.8
UCS2003 K6-101F V0-3B42

232F0
日 72.14
UCS2003 V2-8672

257CC
禾 115.7
UCS2003 V2-7F32

2AE98
爪 87.13
V4-4D5A

All changes are unifiable based on UCV rules.

Character	Code	Normalization Rule
梐	U+23597	1-4
浙	U+23DEF	1-4

飘	U+232F0	1-5
秫	U+257CC	1-5
齛	U+2AE98	3-4

Because stroke counts are already normalized for the most part, only the changes for U+232F0 and U+257CC will result in corresponding Unihan data changes. See details below.

Note. U+232F0, read *bão*, means a "storm", and the source glyph for U+232F0 was originally composed with 凡 (U+322C4), a simplified version of "wind", on right side, as shown on p. 214 of *Bảng tra chữ Nôm*, below:
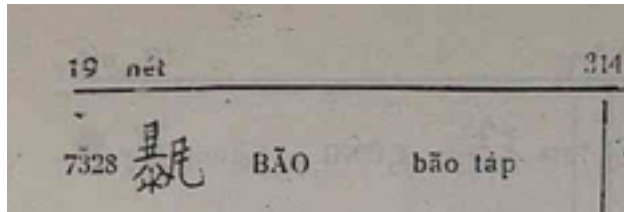
Since the proposed change clarifies the semantic structure of U+232F0, we propose that Unihan radical be changed to 182'' to reflect this. If needed, a secondary radical-stroke value of 72.15 could be added.

Proposed changes

Glyphs:

V-Source	Code	Old glyph	New glyph
V2-7A32	U+23AF2	尋	尋
V3-3570	U+291AC	霥	霥
V2-7938	U+23597	椪	椪
V0-3B42	U+23DEF	浉	浉
V2-8672	U+232F0	飘	飘
V2-7F32	U+257CC	秫	秫
V4-4D5A	U+2AE98	齛	齛
V2-7F33	U+2582D	稣	稣

Unihan data changes:

Code	Tag	old value	new value
U+291AC	kRSUnicode	173.15	173.14
U+291AC	kTotalStrokes	23	22
U+232F0	kRSUnicode	72.14	182''.15
U+232F0	kTotalStrokes	18	19
U+257CC	kRSUnicode	115.7	115.8
U+257CC	kTotalStrokes	12	13