

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG
 Title: IRGN2556R V-Source Glyph and Codes Updates (Revised)
 Source: Vietnam
 Status: National Body Proposal
 Authors: Eiso Chan, Lee Collins, Ngô Trung Việt
 Action: For consideration by the IRG
 Date: 2022-07-11

This document proposes minor changes to 22 Vietnamese Nôm glyphs based on our a desire to more accurately reflect usage, to normalize some of the shapes, and to improve the design of some of the glyphs in the representative NomNaTong font. Four candidates for horizontal extension are proposed to clarify certain distinctions in shape.

1. Horizontal Extensions

Current	TCVN	glyph	details
65B3 斤 69.11 𣎵 𣎵 𣎵 𣎵 𣎵 G5-4E64 HB2-E5C2 T2-5069 J1-4162 K2-3834	VN-065B3	𣎵	Note 6
20302 人 9.10 傘 傘 UCS2003 G4K-12614	V1-4B79	傘	Note 1
29516 頁 181.9 類 類 UCS2003 GHZ-74384.02	VN-29516	類	Note 3
2C13A 木 75.14 𣎵 JK-65332	V2-796F	𣎵	Note 5
6FC2 水 85.13 濂 濂 濂 濂 濂 G0-6525 HB1-BEFC T1-7127 J0-5F32 K0-5632	VN-06FC2	濂	Note 4

2. Glyph Changes

Current						New	Details
5098 人 9.10	傘 G1-4921	傘 HB1-B3CA	傘 T1-5E30	傘 J0-3B31	傘 K0-5F21	傘 V1-4B79	傘 Note 1
6FD3 水 85.13	濂 GE-4033	濂 H-FDD4	濂 T3-5155			濂 V0-3B78	濂 Note 4
826E 艮 138.0	艮 G0-745E	艮 HB1-A6E1	艮 T1-4862	艮 J0-3A31	艮 K0-4A5D	艮 V1-644F	艮 improve design
23E5E 水 85.10	涿 UCS2003	涿 GKX-0639.23	涿 T5-4025			涿 VU-23E5E	涿 improve design
985E 頁 181.10	類 G1-4060	類 HB1-C3FE	類 T1-794B	類 J0-4E60	類 K0-573E	類 V1-6C22	類 Note 5
2087A 力 19.8	勢 UCS2003	勢 V0-3056				勢	勢 Note 2
230B7 斤 69.10	斡 UCS2003	斡 VU-230B7	斡 H-90E4			斡	斡 Note 6
2339E 月 74.12	𠂔 UCS2003	𠂔 V2-7921				𠂔	𠂔 normalize to 月

237C3 木 75.15	𣎵 UCS2003	𣎵 ₂ V2-796F	𣎵	distinguish from U+2C13A
26A5A 舟 137.8	𣎵 ₃ UCS2003	𣎵 ₃ V2-822E	𣎵 ₃	normalize to 舟
26A5B 舟 137.8	𣎵 ₄ UCS2003	𣎵 ₄ V2-822F	𣎵 ₄	normalize to 舟
26A73 舟 137.9	𣎵 ₅ UCS2003	𣎵 ₅ V2-8230	𣎵 ₅	normalize to 舟
26A82 舟 137.10	𣎵 ₆ UCS2003	𣎵 ₆ V2-8232	𣎵 ₆	normalize to 舟
26A90 舟 137.11	𣎵 ₇ UCS2003	𣎵 ₇ V2-8235	𣎵 ₇	normalize to 舟
26AA6 舟 137.12	𣎵 ₈ UCS2003	𣎵 ₈ V2-8236	𣎵 ₈	normalize to 舟
26AA8 舟 137.13	𣎵 ₉ UCS2003	𣎵 ₉ V2-8234	𣎵 ₉	normalize to 舟
27350 虫 142.8	𣎵 ₁₀ UCS2003	𣎵 ₁₀ V2-8331	𣎵 ₁₀	Note 7

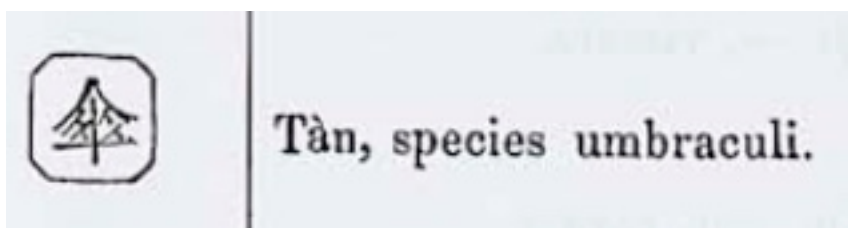
<p>284A3 辵 162.5</p> <p>𨔵 UCS2003 𨔵 GHZ-63825.02 𨔵 T5-2B54 𨔵 VU-284A3</p>	<p>𨔵</p>	<p>improve design</p>
<p>2AB63 手 64.4</p> <p>扌 V4-4873</p>	<p>扌</p>	<p>per IRG52</p>
<p>2C04C 日 72.10</p> <p>𣎵 V4-4A50</p>	<p>𣎵</p>	<p>normalize 芒 > 芒</p>
<p>2BC2E 女 38.7</p> <p>妒 V4-4555</p>	<p>妒</p>	<p>Note 8</p>
<p>26A83 舟 137.10</p> <p>𧨧 UCS2003 𧨧 V2-8233</p>	<p>𧨧</p>	<p>Note 9</p>
<p>2ACD8 木 75.10</p> <p>𣎵 V4-4B3E</p>	<p>𣎵</p>	<p>Note 9</p>
<p>2BF45 手 64.8</p> <p>扌 V4-4931</p>	<p>扌</p>	<p>Note 10</p>

236EF 木 75.12	檻 UCS2003	檻 V2-7951	檻	Normalize to 盜
2C43C 皿 108.14	盜劫 V4-4E7B		盜劫	Note 11
2C43E 皿 108.16	甚盜 V4-4E7C		甚盜	Normalize to 盜
26AD8 色 139.5	餞 UCS2003	餞 V0-422A	餞	Note 12
279F8 言 149.7	謁 GKX-1162.10 謁 VN-279F8	謁 T5-4978	謁	Note 13
20C00 口 30.4	呬 UCS2003	呬 V0-3071	呬	Note 14
2AF6F 疒 104.6	痢 V4-4E56		痢	Note 14

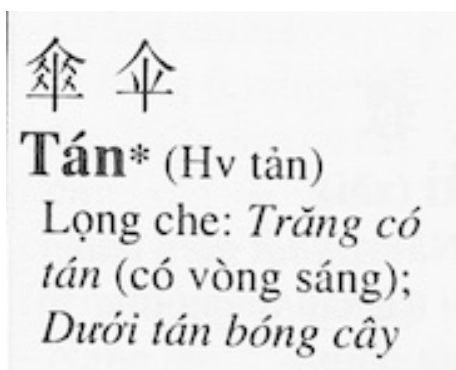
609E 心 61.7 GE-392C	悞	悞	悞	悞	悞	悞	Note 15
23B87 毛 82.7 UCS2003	𦍋	𦍋				𦍋	Note 15
25873 禾 115.11 UCS2003	𪎩	𪎩	𪎩			𪎩	Note 16

Notes

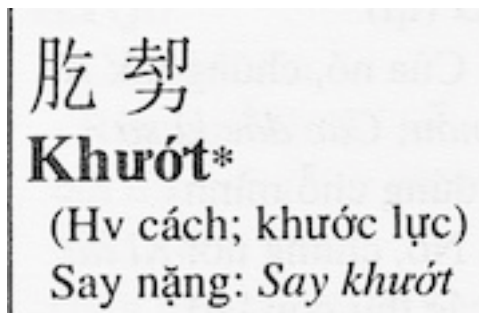
1. U+5098, V1-4B79 Source Taberd has 𠂔人𠂔十𠂔



GĐNHV has U+20302 𠂔人𠂔𠂔十

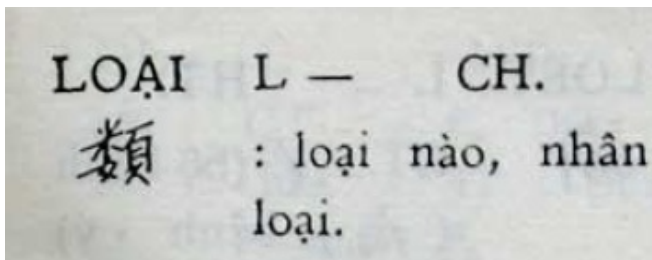


2. U+2087A, V0-3056, *khưót* in GĐNVH uses the phonetic 却 *khưóc*.



3. U+29516, VN-29516

V1-6C22 is currently mapped to U+985E. The original source for V1-6C22, Vũ Văn Kính's *Tự Điển Chữ Nôm*, was actually a different character, U+29517 = VN-29517:



The form with 犬 is more common, but 大 is also found, as for example, in Takeuchi Yonosuke, *Tự Điển Chữ Nôm*

loài	類 : loạiの変音, 同義。	loài cây. 類核。(植物)
		loài vật. 類物。(動物)
		loài người. 類俤。(人類)

loại	類 : 音(loại)の借用。	loại ra. 類器。
------	-----------------	--------------

Given the lack of an authoritative source for V1-6C22, we will keep V1-6C22 mapped to U+985E, change the glyph to use the more common form, and horizontally extend VN-29516.

4. Both U+6FD3 and U+6FC2 are found in Vietnamese: V0-3B78 is in Vũ Văn Kính's *Tự Điển Chữ Nôm*. The new shape matches TCVN V0-3B78 so no change of source.

NHÈM L — HT.
 濂 CT- 𣶒 (bộ thủy
 (水): nước → ý)
 +(liêm → âm) 廉
 : lèm nhèm, nhập
 nhèm.

VN-06FC2 is found in GĐNHV

洊 濂
Nhèm* (thủy nhiễm)
 (thủy liêm)
 - Có vết: *Nhèm mực*
 - Không rõ ràng: *Số*
sách nhập nhèm;
Chữ viết lèm nhèm

5. U+2C13A, V2-796F, GĐNHV

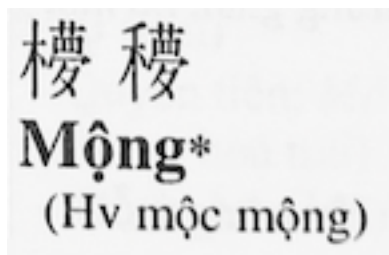
As seen below, the original evidence cited for U+237C3 V2-796F is clearly not 𣶒木𣶒苗一歹:

mộng	樓	2-37C3	V2-796F	0561.111	L295.20
------	---	--------	---------	----------	---------

The source cited, L295.20 is character no. 20 on p. 295 of *Bảng Tra Chữ Nôm*:

6864 樓 MỘNG mọc mộng,
 mộng cửa

We see the same form in another source, *Giúp đọc Nôm và Hán Việt*:

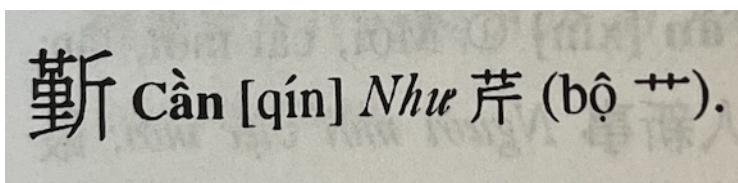


We changed the glyph for V2-796F to 𣏟木夢 in v3.9 of the NomNaTong font, before we were aware that U+2C13A was available in Extension E. This is what's shown in KCHN.

mộng	夢	夢	懵	梦
	U+05922	U+05923	U+061F5	U+068A6
	樓	樓	瞢	樓
	U+237C3	V+6036F	U+077A2	V+6042F
	𦵒	𦵒	𦵒	𦵒
	U+04445	U+26D0B	U+28B92	U+09E0F

We cannot find examples of 𣏟木𣏟𣏟𣏟, but we believe that having the shape of the encoded character more closely match the shape of the glyph, and not be easily confused with another character is long term advantage to type designers that outweighs any short-term disadvantage.

6. U+65B3, VN-065B3, *Tự điển Hán Việt Hán ngữ cổ đại và hiện đại*.

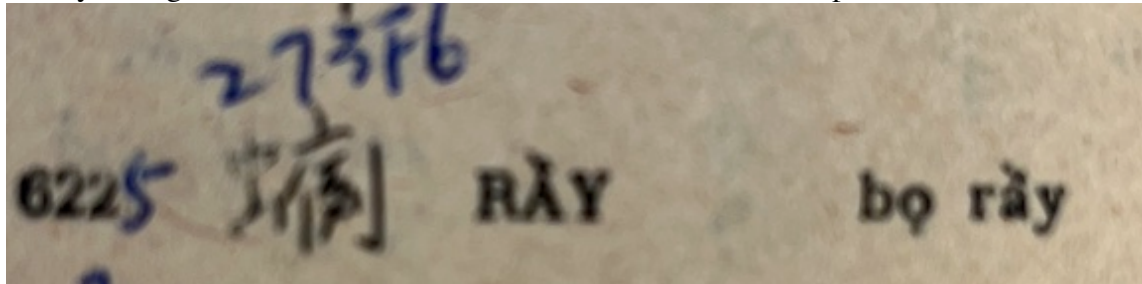


In *Đại Việt Sử Ký Toàn Thư* U+ 230B7 is used for the name Đỗ Cậ

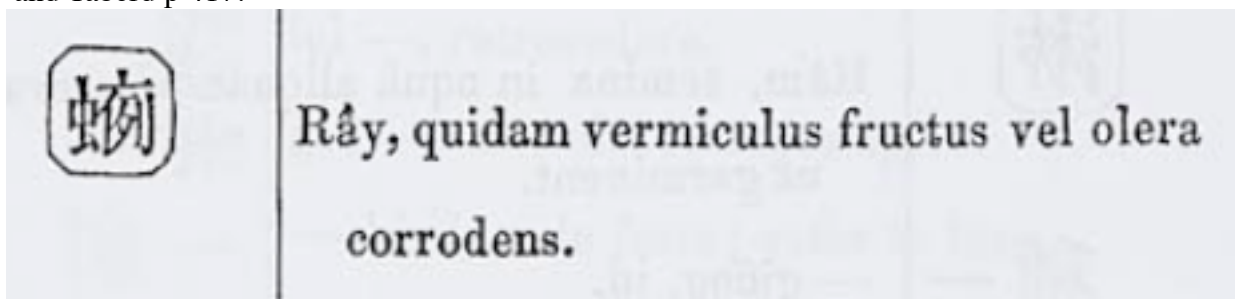


7. U+27350.

While 列 and 例 are not unifiable, we need to change U+27350 because the phonetic 列 (*liệt*) is clearly wrong and does not match the sources, which are BTCN p. 271



and Taberd p 417.



The form found in the UCS is the only example we have of a glyph with the 𧈧 列 component. The others with similar phonetic all contain 𧈧 (*lời*). Prof. Nguyễn Quang H ồng believes this glyph to be derived from a combination of 麻 *ma* and 例 *lệ* (𧈧 *lời*), which represented an initial **ml* that existed in an earlier form of Vietnamese spoken when the characters were first invented. This change is essential from the viewpoint of historical linguistics.

8. U+2BC2E

The source for V4-4555 is p. 515 of the *Góp phần nghiên cứu văn hoá Việt Nam*, a reference work on the Tày language:

TÀY			YẾU TỐ NGŨ ÂM			YẾU TỐ NGŨ NGHĨA		
lêng	𧈧	nuôi	lệnh	𧈧	lêng	thực	𧈧	ăn
lính	𧈧	con khỉ	lính	𧈧	lính	khuyến	𧈧	con chó
lúa	𧈧	con dâu	lô	𧈧	lou	nữ	𧈧	con gái

These 2 characters wrongly use 冢 (SV: mông) as the phonetic, where it should be 冢 (SV:

The current form of V4-4931 捌, 𢆶去𠂔力𠂔, is unusual. Prof. H ồng in *Từ Điển Chữ Nôm*

CƯỚP 勦 HT Trợ 助 + Kiếp 劫 của, vợ -
chồng, Trộm -

盜劫 HT Đạo 盜 + Kiếp 劫

Note: the stroke counts remain the same.

甚盜 HT Thộm 甚 + Đạo 盜

12. U+26AD8

The right side of 𪗇 V0-422A is not a simplification from 發. The source, *Tự điển chữ Nôm* by Vũ Văn Kính, shows right side of 拔, which we have normalized as 𢇛.

BỘT L- HT.
𪗇 CT- 色 (sắc :
sắc→ý)+(1/2 chữ
bạt (拔) →
âm). 𢇛
: Bột màu : phai.

13 U+279F8

Normalize 𪗇 VN-279F8 to 𪗇, which is more common in Vietnamese. We are separately proposing normalization of 𪗇 VN-F1BF3 and 𪗇 VN-F0B89 in WS 2021.

14. U+20C00 and 2AF6F

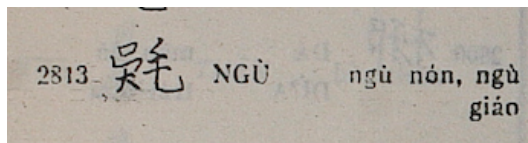
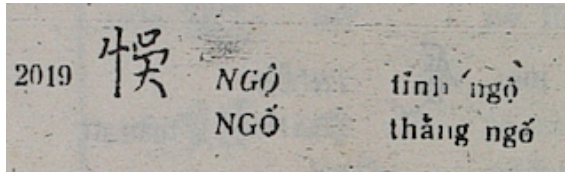
The normal forms 𪗇 V0-3071 and 𪗇 V4-4E56 should use 𪗇, as shown in the original source, *Tự điển chữ Nôm* by Vũ Văn Kính,

DIỀU L- HT.
𪗇 CT- 𪗇 (khẩu :
miệng→ý + (điều
→ âm) 𪗇
: bêu điều

We are separately proposing normalization of 𐄀 VN-F1629 in WS 2021.

15. U+609E and U+23B87.

The current shape of the 𠂔 element in 𠂔 V2-8C25 and 𠂔 V2-7A35 is unusual. The source, *Bảng Tra Chữ Nôm*, has the normal forms.



16. 𪔐 U+25873 VN-25873

The source, Taberd, has 𪚩𪚪𪚫, not 𪚩𪚪𪚫𪚬𪚭𪚮𪚯𪚰𪚱𪚲.

<div style="border: 1px solid black; padding: 5px; display: inline-block; margin-bottom: 10px;">𪚩</div> <p>— 憾</p> <p>— 𪚩</p> <p>— 𪚩</p>	<p>Ghê, dentium stupor.</p> <p>gồm —, maximoperè: valdè, (de re bonâ et malâ.)</p> <p>chỉn —, idem, (sed de re bonâ tantùm.)</p> <p>— thay, id.</p> <p>— răng, dentes hebetari.</p>
--	---

Data changes:

		old	new
U+5098	kIRG_VSource	V1-4B79	VN-05098
U+65B3	kIRG_VSource	—	VN-065B3
U+6FC2	kIRG_VSource	—	VN-06FC2
U+20302	kIRG_VSource	—	V1-4B79
U+237C3	kIRG_VSource	V2-796F	VN-F20C2
U+29516	kIRG_VSource	—	VN-29516
U+2C13A	kIRG_VSource	—	V2-796F
U+2087A	kRSUnicode	19.8	19.7
U+236EF	kRSUnicode	75.12	75.11
U+236EF	kTotalStrokes	16	15
U+27350	kRSUnicode	142.8	142.10
U+2AB63	kRSUnicode	64.4	64.5
U+2AB63	kTotalStrokes	7	8
U+2BC2E	kRSUnicode	38.7	38.8
U+2BC2E	kTotalStrokes	9	10
U+2C04C	kRSUnicode	72.10	72.11
U+2C04C	kTotalStrokes	14	15
U+2C43E	kRSUnicode	108.16	108.15
U+2C43E	kTotalStrokes	21	20