

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG
 Title: IRGN2556R2 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: 2022-10-20

This document proposes minor changes to 33 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. Five 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
2585E 禾 115.11 𣎵 𣎵 UCS2003 GKX-0858.17	VN-2585E	𣎵	Note 16

2. Glyph Changes

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

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

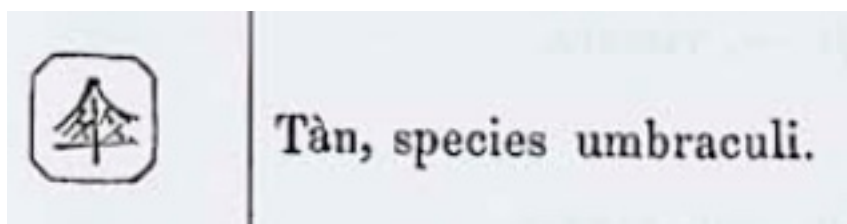
<p>284A3 𨔵 162.5</p> <p>𨔵 UCS2003 𨔵 GHZ-63825.02 𨔵 T5-2B54 𨔵 VU-284A3</p>	<p>𨔵</p>	<p>improve design</p>
<p>2AB63 手 64.4</p> <p>𢮗 V4-4873</p>	<p>𢮗</p>	<p>per IRG52</p>
<p>2C04C 日 72.10</p> <p>𣎵 V4-4A50</p>	<p>𣎵</p>	<p>normalize 芒 > 芒</p>
<p>2BC2E 女 38.7</p> <p>妒 V4-4555</p>	<p>妒</p>	<p>Note 8</p>
<p>26A83 舟 137.10</p> <p>𩇛 UCS2003 𩇛 V2-8233</p>	<p>𩇛</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>𢮗 V4-4931</p>	<p>𢮗</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	甚次 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 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

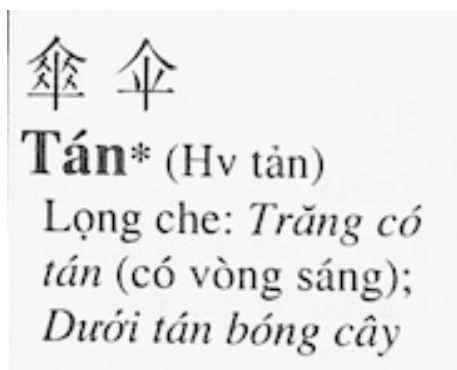
609E 心 61.7 悞 GE-392C 悞 H-FC6B 悞 T3-3127 悞 J13-7452 悞 K2-324B 悞 V2-8C25	悞	Note 15
23B87 毛 82.7 𦍋 UCS2003 𦍋 V2-7A35	𦍋	Note 15

Notes

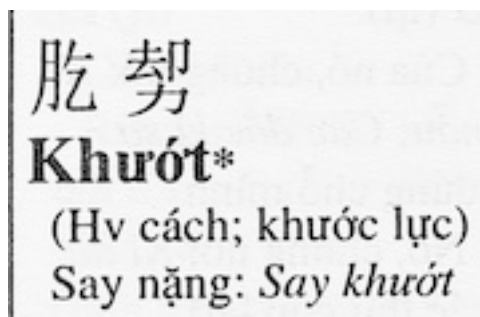
1. U+5098, V1-4B79 Source Taberd has 𠂔人𠂔十𦍋



- GĐNHV has U+20302 𠂔人𠂔𦍋十

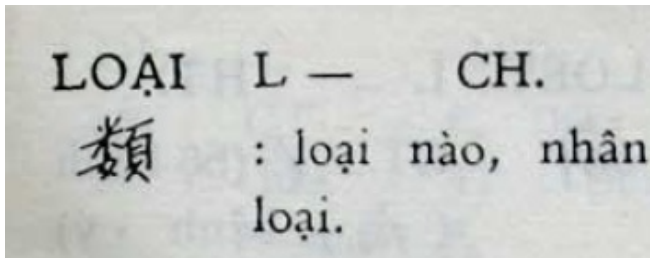


2. U+2087A, V0-3056, *khurót* in GĐNVH uses the phonetic 却 *khuróc*.



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:

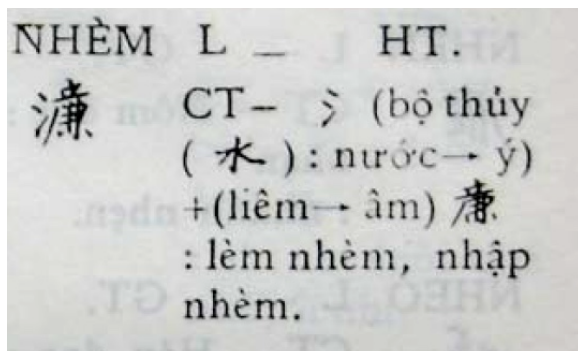


The form with 犬 is more common, but 大 is also found, as for example, in Takeuchi Yonosuke, *Tự Điển Chữ Nôm*

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



VN-06FC2 is found in GĐNHV

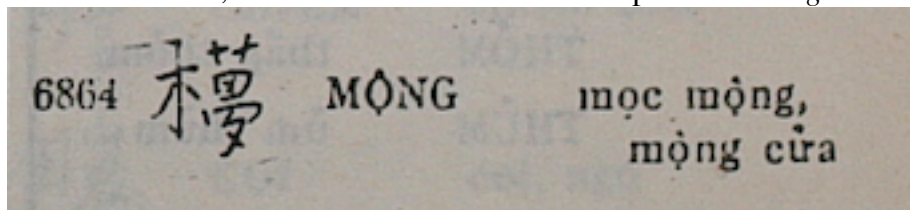


5. U+2C13A, V2-796F, GĐNHV

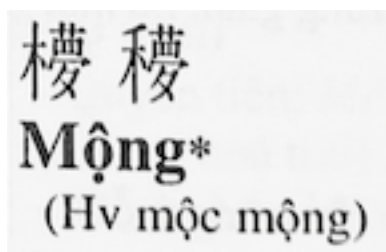
As seen below, the original evidence cited for U+237C3 V2-796F is clearly not 𣎵木𣎵苗𣎵歹:

mộng	𣎵	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*:



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

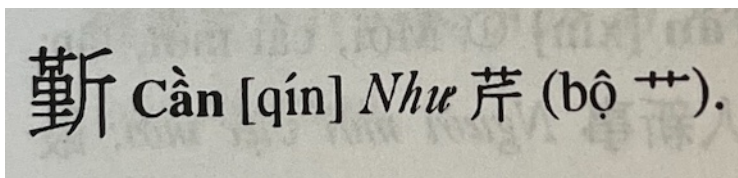


We changed the glyph for V2-796F to 𣎵木夢 in v3.9 of the NomNaTong font, before we were aware that U+2C13A was available in Extension E. This is what's shown in KCHN.

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

We cannot find examples of 𣎵木𣎵苗一𠂔, 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*.

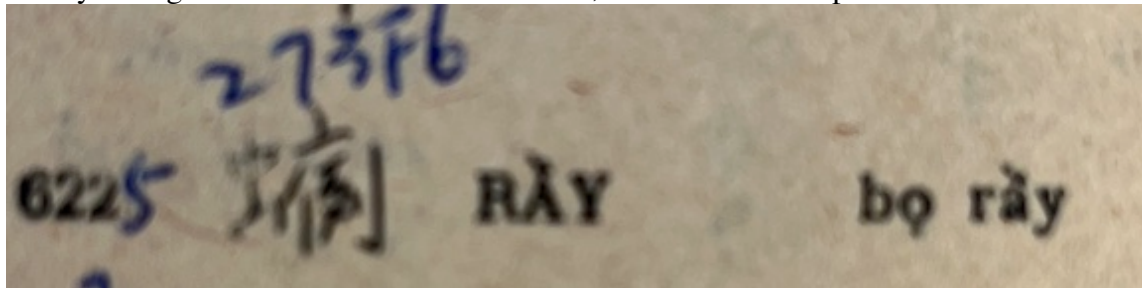


In *Đại Việt Sử Ký Toàn Thư* U+ 230B7 is used for the name Đỗ Cậ

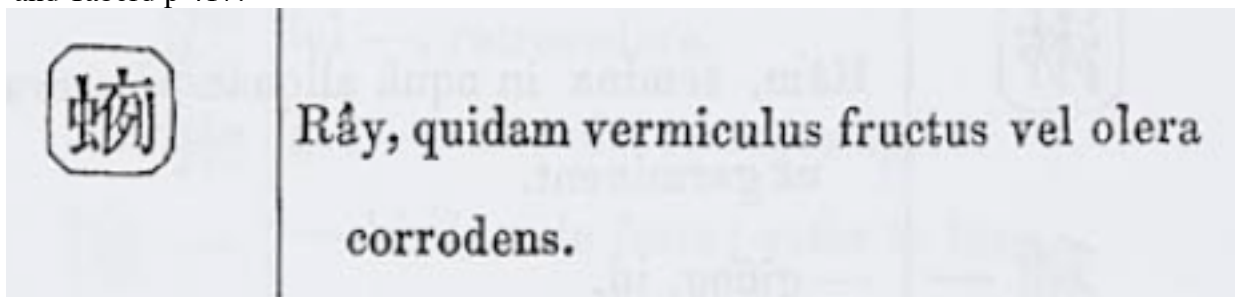


7. U+27350.

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 例 (*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 TỔ NGỮ ÂM			YẾU TỔ NGỮ NGHĨA		
lêng	𧈧	nuôi	lệnh	令	ling	thực	食	ăn
lính	𧈧	con khỉ	lính	灵	ling	khuyến	𧈧	con chó
lúa	𧈧	con dâu	lô	芦	lou	nữ	女	con gái

The analysis of this character, read *lúa* and meaning *daughter-in-law*, shows that it is composed of the phonetic element 芦, read *lô* in Vietnamese and *lou* in Chinese (eq. to Pinyin *lu*), and the

9. U+26A83 and U+2ACD

𦨭 HT Chu舟 + Trùng冢

KHÉP # {F2: thủ 扌 𢵿 kiếp} . ◎ Hãm vào, đẩy vào chốn nào đó.
 𢵿 U+2BF45 丕生畚𢵿咍制 鞞麻吏搯𠂔尼空錢 Trời sinh ra kiếp hay
 chơi. Sao mà lại khép vào nơi không tiền (Nam lục, 26b). Ø CƯỚP
 KHÉP.

The normal form of 𪔐 V4-4E7B is 盜, with ice, not water. Also, the source *Đại Tự Điển Chữ Nôm*, has 𪔐, not 劫:

CUỐP 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

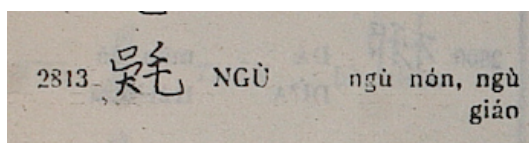
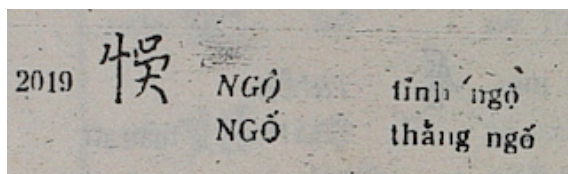
The normal forms 𪗇 V0-3071 and 𪗇 V4-4E56 should use 𪗇, as shown in the original source, *Tự điển chữ Nôm* by Vũ Văn Kính,

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

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

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

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



16. 替 U+2585E VN-2585E

The source, Taberd, has 𪛗𪛗𪛗, not 替, 𪛗𪛗𪛗.

替	Ghê, dentium stupor.
一 憾	gồm —, maximoperè : valdè, (de re bonâ et malâ.)
一 𪛗	chỉn —, idem, (sed de re bonâ tantùm.)
台 一	— thay, id.
𪛗 一	— răng, dentes hebetari.

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+2585E	kIRG_VSource	—	VN-2585E
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