

PROPOSAL

TO STREAMLINE U+96B9 隹 AS A COMPONENT IN V-SOURCE REPRESENTATIVE GLYPHS

Individual Contribution

For consideration by UTC

Alexander Zapryagaev

18 August 2022

rudetection@gmail.com

The glyphs appearing in the Unicode code charts serve a double purpose. Beside giving a convenient listing of the characters encoded, they also provide the information for font developers about at least one correct implementation of the characters. The representative glyphs for the CJK Unified Ideographs do an additional duty: representing the difference between various national standards.

However, the fonts used for these representative glyphs in not always consistent, leading to conflict between the theoretical correct regional form and the glyph seen in the tables as well as confusion for the potential font makers, unsure whether the peculiarity they see is to be implemented. A well-documented example of such a discrepancy in the G-Source, reported by Andrew West, is the form of 𠂇 radical in characters sourced to GZJW (Yīnzhōu Jīnwén Jíchéng Yīndé 殷周金文集成引得), which is different compared to the form used in the remaining G-Source characters.

This document proposes to eliminate one of the possible sources of confusion by streamlining the presentation of the component U+96B9 隹 (U+2FAB 隹 KANGXI RADICAL SHORT TAILED BIRD) in the V-Source glyphs (the font Nom Na Tong Regular). The orthography of U+96B9 隹 is well-known and consistent in most of the localities. The form used in the Kangxi Dictionary, which we will denote as 隹 “parallel”, is normal for Japanese and Korean

localities, while the form 隹 “opposite” occurs in the PRC, as well as Taiwan and Hong Kong. The study of the representative glyphs in the Unicode code tables confirms these findings.

However, the situation with the Vietnamese form of the same element is not so clear-cut. Currently, there is no obvious prevalence of one form over the other; even within the derivative of the same first-level glyph, such as in the family of 雍, there may be discrepancies – compare, for example, 雍 and 擁.

There is no apparent connection with the particular source, despite the newer the source (V3, V4, VU) the more prevalent the “parallel” form 隹 becomes (but U+2C6CE 離 is sourced to V4 and has “opposite”).

Numerically, there are much more forms containing the “parallel” form. No pattern in Hán vs. Nôm usage could be found.

Such issues complicate the process of determining which details in glyphs are actually relevant local forms. For example, all V-Source glyphs consistently show 言 with vertical first stroke, or render the 讠 component at the left in a small form, shifted to the top¹. Thus, it would be fair to assume that both of these are, indeed, national features.

On the other hand, the form of 隹 happens to be not. Informal discussions with Nôm users did not procure an actual user preference. The difference in the forms of representative glyphs seems to be a mere encoding accident. Hence, when looking at a particular character with a glyph, it becomes more problematic to determine whether this is a local feature that should be implemented in a font no matter what.

In fact, no other font purporting to represent the Chữ Hán and the Chữ Nôm characters even attempts to follow this distribution. For example, *Han Nom A/B* provides mostly “opposite” glyphs barring some exceptions; on the other hand, *Han Nom Gothic* overwhelmingly produces the “parallel” form, also with some (rare) exceptions.

I propose the Vietnamese representatives to correct 38 V-Source representative glyphs (and 41 glyphs of potential horizontal extensions) in the chart in order to bring all instances of U+96B9 隹 component to “parallel” form 隹.

¹ This is due to a lot, probably most, examples of 讠 讠 being occurrences of diacritic *khẩu*, unified to 讠 in Unicode, not components of Hán characters (see [L2/18-318](#), p. 4).

APPENDIX I: 218 CHARACTERS WITH COMPONENT 隹 HAVING V-SOURCE

38 are “opposite” and require change:

Code	Glyph	V-Source	Code	Glyph	V-Source
U+5095	隹	V2-8A40	U+96C5	雅	V1-6B43
U+51C6	淮	V1-4C63	U+96C7	雇	V2-912B
U+552F	唯	V1-4F42	U+96C9	雉	V1-6B45
U+5806	堆	V1-5075	U+96CD	雍	V1-6B47
U+60DF	惟	V2-8C28	U+96D6	雖	V1-6B48
U+63A8	推	V1-5743	U+96DC	雜	V1-6B4A
U+6409	推	V2-8C53	U+96E2	離	V1-6B4B
U+690E	椎	V1-5A38	U+96E3	難	V1-6B4C
U+6DEE	淮	V1-5C64	U+9A05	騅	V1-6C47
U+731A	雅	V2-8D74	U+9DB4	鶴	V1-6C79
U+7515	甕	V1-5F43	U+9E0C	鸛	V0-4858
U+78BA	確	V1-606E	U+20ED3	喙	V2-714B
U+7A1A	稚	V2-8E43	U+210DC	喙	V2-7257
U+81C3	臙	V2-8F32	U+21457	堆	V0-346A
U+8AB0	誰	V2-9027	U+243E5	燿	V2-7B6D
U+9032	進	V1-6947	U+28192	躡	V2-8475
U+9310	錐	V1-6A4F	U+281F9	躡	V2-8535
U+96BE	难	V2-6F71	U+29833	首隹	V2-8729
U+96C4	雄	V1-6B42	U+2C6CE	肉隹	V4-5230

180 are “parallel” and okay:

Code	Glyph	V-Source	Code	Glyph	V-Source
U+50AC	催	V1-4B7B	U+22975	懣	VN-22975
U+52F8	勸	V1-4D51	U+22984	懣	V2-7639
U+5684	嘍	V1-5028	U+22A64	懣	V2-763C
U+56AF	嚯	V1-5032	U+22C9C	搯	VN-22C9C
U+56C4	嚙	V0-337D	U+22D78	搯	V0-3865
U+56C9	囉	V1-5036	U+22DA5	搯	V2-7760
U+596A	奪	V1-5158	U+22DBE	搯	V0-3869
U+596E	奮	V1-515A	U+22E42	攜	VN-22E42
U+5D14	崔	V1-534B	U+22E9C	攢	V2-782E
U+615B	催	V1-5572	U+23125	競	V2-7840
U+6194	憔	V1-5623	U+23272	瞋	V0-3955
U+61C9	應	V1-562D	U+2328C	瞋	V2-785D
U+622A	截	V1-563D	U+23309	鬪	VN-23309
U+64C1	擁	V2-8C67	U+2366F	樵	V2-7957
U+64E2	擢	V2-8C70	U+236DC	樵	VN-236DC
U+6509	擢	V2-8C73	U+23F2D	灘	V0-3B62
U+651C	攜	V1-5773	U+23F32	灘	V3-3353
U+651E	擺	V0-3938	U+240DF	漣	V0-3C28
U+6524	攤	V0-3937	U+24137	灑	VN-24137
U+66B9	暹	V1-586C	U+24350	推	VN-24350
U+69EF	權	V1-5A56	U+244CC	爓	VN-244CC
U+6A35	樵	V2-8D3D	U+244D3	爓	VN-244D3
U+6AC2	權	V1-5A79	U+244DB	爓	V3-337E
U+6B0A	權	V1-5B22	U+245EF	耀	VN-245EF
U+6B61	歡	V1-5B31	U+24ADA	璫	VN-24ADA
U+6E96	準	V1-5D26	U+24D0F	巽	V2-7D23

U+6F3C	漼	V1-5D44	U+24EAE	癩	V2-7D47
U+6F50	漼	V1-5D45	U+2502A	癩	V2-7D67
U+6F57	漼	V1-5D47	U+25296	睺	V0-3E70
U+71A3	燿	V0-3C63	U+256ED	襪	V0-4769
U+7372	獲	V1-5E76	U+256F2	襪	VN-256F2
U+7670	癩	V1-5F7A	U+258FD	襪	V3-3432
U+77BF	瞿	V2-8E38	U+25D1B	籊	V2-7F6E
U+77CD	瞿	V2-8E39	U+25D22	籊	V2-7F74
U+7A6B	穫	V1-6148	U+25DAE	籊	V2-7F7B
U+7E99	羅	V0-4134	U+25DE7	籊	V3-3654
U+7F85	羅	V1-6344	U+25E01	籊	V2-8026
U+8000	耀	V2-6E75	U+26349	縞	V0-474B
U+8549	蕉	V1-6554	U+265E0	瞧	VN-265E0
U+8638	蘸	V2-8F5C	U+267D7	臙	V2-816A
U+863A	蘸	V1-656A	U+267FF	臙	V2-8162
U+863F	蘿	V0-426F	U+26864	臙	V0-4178
U+8667	虧	V1-6571	U+268CE	臙	V2-8224
U+8862	衢	V2-8F71	U+269C3	臙	V2-822A
U+89C0	觀	V1-666D	U+26FA3	摧	V2-8262
U+8B77	護	V1-6763	U+270F7	蘆	V2-8274
U+8B8E	讎	V1-6767	U+270FD	蘆	V2-8272
U+8E8D	躍	V1-6862	U+27110	夔	VN-27110
U+908F	邏	V1-696A	U+2743B	蚩	V2-833E
U+944A	鑊	V1-6A68	U+2753F	蚩	V2-8353
U+947C	鑊	V0-4660	U+277A4	襪	V0-4369
U+96BB	隻	V1-6B40	U+277DD	襪	V2-836B
U+96C0	雀	V1-6B41	U+27939	攜	V2-8370
U+96C6	集	V1-6B44	U+2811F	墮	V0-4534

U+96CB	雋	V1-6B46	U+2813A	踫	V2-8469
U+96D5	雕	V2-912D	U+28920	糴	VN-28920
U+96D9	雙	V1-6B49	U+28B09	錐	V2-857A
U+970D	霍	V1-6B53	U+28B36	錐	VN-28B36
U+9867	顧	V1-6C23	U+28B91	錐	V2-8627
U+9DF9	鷹	V1-6C7E	U+28FFD	錐	VN-28FFD
U+9E1C	鸚	V0-4859	U+29033	雉	V2-8645
U+35F1	噍	V2-8924	U+29066	馥	V2-8646
U+457C	蕪	V2-8975	U+29071	馥	V0-486E
U+49F9	雁	V2-912C	U+2915A	霽	V3-356C
U+20371	雋	VN-20371	U+291BE	霽	V2-865B
U+20438	雉	V0-456A	U+29353	鞞	V2-8666
U+206A2	黜	V0-3F6E	U+29E79	雙	V0-4836
U+20F32	唯	V0-3322	U+29F67	雙	VN-29F67
U+20F92	唯	V2-7170	U+2A165	雙	V3-363C
U+20FF2	𪛗	V2-716D	U+2A167	鳩	V2-8854
U+20FF6	𪛗	V3-3058	U+2A192	鷹	VN-2A192
U+21030	唯	V2-7175	U+2A23E	鷹	VN-2A23E
U+210FC	噍	V2-7253	U+2A771	催	V4-4145
U+21110	噍	V0-3374	U+2A7E6	雅	V4-4249
U+2112A	唯	V0-337A	U+2A9FE	雅	V4-4639
U+21167	噍	V0-3422	U+2AA05	雅	V4-4655
U+21173	勸	V2-726B	U+2ABF0	雅	V4-4A25
U+21196	噍	V0-3428	U+2AE70	雅	V4-4D45
U+21197	噍	V3-3063	U+2AE87	雅	V4-4D52
U+2119A	噍	V2-7273	U+2AEC9	雅	V4-4D77
U+214CB	堆	V3-3443	U+2B84D	雅	V4-404E
U+21503	堆	V2-7352	U+2B904	雅	V4-414C

U+2165C	𨮑	VN-2165C	U+2B911	𨮑	V4-4153
U+2169D	𨮑	V3-3122	U+2BBA5	𨮑	V4-4472
U+21CEA	𨮑	V0-4623	U+2BC5E	𨮑	V4-456E
U+21ECE	𨮑	VN-21ECE	U+2C23C	𨮑	V4-4C5B
U+21F9C	𨮑	VN-21F9C	U+2C774	𨮑	V4-5327
U+21FC7	𨮑	VN-21FC7	U+2C96C	𨮑	V4-545D
U+222EC	𨮑	V2-7546	U+2CBF0	𨮑	V4-5659
U+2296A	𨮑	V0-3669	U+2CDE9	𨮑	V4-5823

APPENDIX 2: 178 CHARACTERS WITH COMPONENT 隹 NOT HAVING V-SOURCE BUT PRESENT IN NOM NA TONG AND THUS QUALIFYING FOR A HORIZONTAL EXTENSION

41 are “opposite” and require change:

Code	Glyph	Code	Glyph
U+5020	隹	U+96BA	隹
U+58C5	壅	U+96CC	雌
U+5A4E	雉	U+96CE	雉
U+5E37	帷	U+96CF	雉
U+69B7	榷	U+96D2	雉
U+6FAD	滌	U+96DB	雛
U+737E	獯	U+96DD	雛
U+74D8	瓘	U+96DE	雞
U+7655	癱	U+9954	饗
U+7762	睢	U+9A69	驩
U+7893	確	U+9DBD	鷓
U+7DAD	維	U+9E64	鶴
U+7FDF	翟	U+3FC8	癱
U+813D	臚	U+49FA	雄
U+8579	蕪	U+2430D	燿
U+873C	雌	U+26E4F	蕪
U+8B99	謹	U+28F88	雇
U+8C01	谁	U+2B534	閨
U+8C9B	獯	U+2D332	雛
U+9180	醜	U+2E977	雉
U+9525	錐		

137 are “parallel” and okay:

Code	Glyph	Code	Glyph
U+50EC	焦	U+853A	蔺
U+50F1	僱	U+8599	薙
U+5101	僑	U+85FA	蔺
U+5138	僱	U+85FF	藿
U+513A	僱	U+8816	蠖
U+51D6	準	U+8837	蠖
U+5281	劓	U+883C	蠖
U+52E7	勸	U+894D	襪
U+532F	匯	U+8B59	譙
U+552E	售	U+8B90	讐
U+55FA	噉	U+8C2F	譙
U+564D	噉	U+8D0B	質
U+5D95	嶠	U+8D5D	質
U+5DB2	嶠	U+8DAF	趨
U+5DC2	嶠	U+8E99	躡
U+5FC2	翟	U+91AE	醮
U+61FC	懼	U+93B8	鏹
U+61FD	懼	U+93D9	鏹
U+6233	戮	U+942B	鏹
U+643A	携	U+9475	鏹
U+644A	摊	U+9481	鏹
U+6467	摧	U+954C	鏹
U+64D3	攤	U+956C	鏹
U+64ED	獲	U+95B5	閩
U+6521	攤	U+96B9	隹
U+652B	攫	U+96BC	隹

U+66DC	曜	U+96BD	隼
U+69AB	樺	U+96C1	雁
U+69DC	携	U+96D0	虐
U+6A29	榷	U+96D4	雥
U+6A87	檣	U+96DA	萑
U+6B0B	櫂	U+96E5	隳
U+6B0F	櫂	U+9866	鰓
U+6B14	櫻	U+9874	顴
U+6C0D	毳	U+98A7	顴
U+6ED9	滙	U+9DE6	鷓
U+6EE9	滩	U+9E1B	鷓
U+6FC9	滩	U+9E6A	鷓
U+6FEF	濯	U+9E70	鷹
U+7048	濯	U+9E73	鷓
U+704C	灌	U+9FA8	匪
U+7055	灘	U+3560	獲
U+7058	灘	U+3617	嚳
U+7126	焦	U+3A26	獲
U+71FF	燿	U+3A57	携
U+7383	獲	U+3A73	獲
U+7480	璀	U+3A74	擢
U+762B	癱	U+3C0C	權
U+766F	癰	U+3C19	攤
U+7671	癰	U+425C	籊
U+77A7	瞧	U+4326	繹
U+77C5	曜	U+447E	艘
U+7901	礁	U+48E4	鄲
U+7C4A	籊	U+4986	饗

U+7C6C	籬	U+4A07	雙
U+7C6E	籬	U+4CFD	鶴
U+7CF4	籹	U+21494	墾
U+7CF6	籹	U+24705	穽
U+7F50	罐	U+26F94	舊
U+7F79	穽	U+28F8F	隻
U+81B2	膳	U+292D8	鞋
U+81BA	膺	U+2AE6F	燄
U+81DB	臙	U+2BACA	噉
U+81DE	臙	U+2C465	睽
U+81E1	齏	U+2CBF5	瞿
U+820A	舊	U+2D13F	勸
U+8411	萑	U+2DB08	權
U+84A6	萑	U+2DFF3	驪
U+84F7	萑		