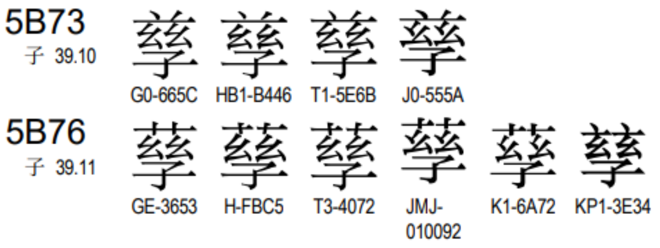
L2/25-013

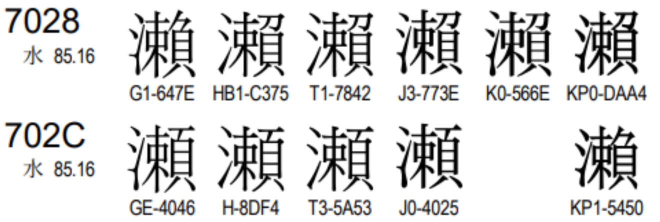
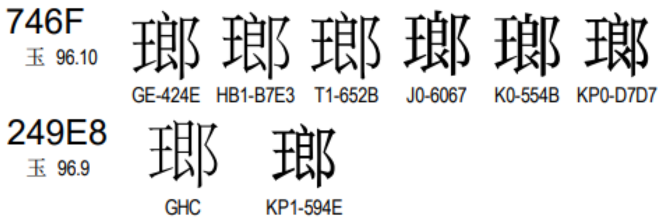
Source:	CheonHyeong Sim (沈天珩) *as the main stakeholder of KP–source glyphs*
Title:	UTN #50 Additions
Date:	2025–01–22
Action:	To be considered by CJK & Unihan WG and UTC

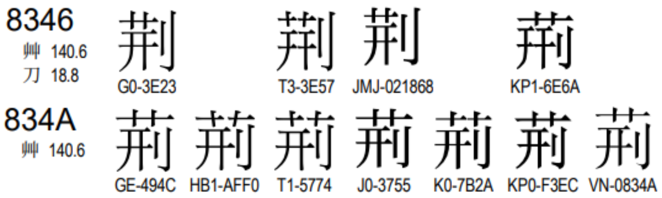
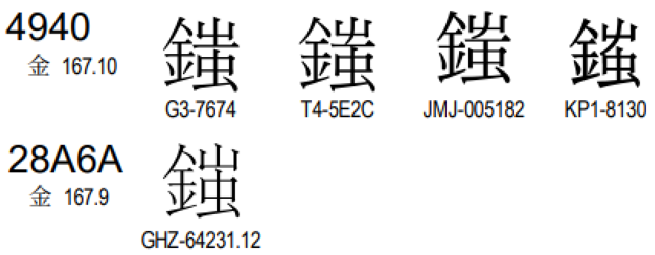
Based on the public review issues reported by Judith Chen, Andrew West (see section 12 of L2/24–067) and Eiso Chan (see section 4 of L2/25–009), this document proposes to add some additional information to UTN #50. (FYI, the document reconstructed KPS 10721 was updated to the 6th version on 2025–01–17)

According to the 勘误 section and the 附录一 section in the reconstructed KPS 10721 document, there are many suboptimal mappings for KP–source characters. A small part of them were reported by Judith Chen and Andrew West as mentioned in the first paragraph of this document. Here, I will list the table with some additional comments, and propose to add the table to the excel file of UTN #50 as Sheet 2. The excel version of this table (without the comments) is attached in this proposal, and I recommend the CJK & Unihan WG to review all the characters with the comments to decide whether they should be added to UTN #50 and whether additional comments are needed for UTN #50.

Table 1. Suboptimal Mappings

	KP1 codepoint	current Unicode	better Unicode	comments
丗	KP1–3413	4E17	2000D	*reported by Judith*
𡟮	KP1–3DB4	217EE	5A90	
孶	KP1–3E34	5B76	5B73	 5B73 子 39.10 G0-665C HB1-B446 T1-5E6B J0-555A 5B76 子 39.11 GE-3653 H-FBC5 T3-4072 JMJ-010092 K1-6A72 KP1-3E34 Glyph different to both. Sim personally thinks 5B73 would be closer.
尒	KP1–3EC8	5C12	21B57	
尵	KP1–3EE4	5C35	2BD52	

屮	KP1−3F1C	5C6E	4DB9	Glyph closer to 4DB9. Hard to confirm the abstract shape semantically.
愸	KP1−439A	6138		Should be disunified. Ref. IRGN2782
𢥻	KP1−44BD	2297B	2D7E7	Cognate but not UCV.
敻	KP1−4850	657B	22FCC	
暅	KP1−49AC	6685	23236	
曚	KP1−4A00	66DA	232E1	
㯞	KP1−4D4C	3BDE	23693	*reported by Andrew*
𣧷	KP1−4F56	239F7	2C194	
瀬	KP1−5450	702C		 7028 水 85.16 G1-647E HB1-C375 T1-7842 J3-773E K0-566E KP0-DAA4 702C 水 85.16 GE-4046 H-8DF4 T3-5A53 J0-4025 KP1-5450 Glyph identical to 7028. Yet 7028 occupied by KP0−DAA4.
𤑔	KP1−5642	24454	2C2D6	Cognate but not UCV.
爋	KP1−5653	720B	24455	*reported by Andrew*
爢	KP1−5662	7222	3E0F	Cognate but NUCV.
獡	KP1−5863	7361	2486F	
𤧨	KP1−594E	249E8		 746F 玉 96.10 GE-424E HB1-B7E3 T1-652B J0-6067 K0-554B KP0-D7D7 249E8 玉 96.9 GHC KP1-594E Glyph identical to 746F. Yet 746F occupied by KP0−D7D7.
𤫊	KP1−59E2	24ACA	2919C	

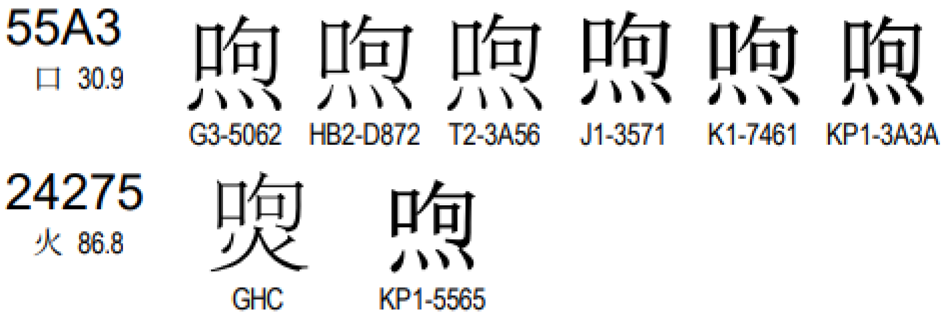
𤬾	KP1−5A33	24B3E	2C3A8	
𤱧	KP1−5AE3	24C67	2DED0	
疉	KP1−5B21	7589	24D01	
盏	KP1−5D4F	76CF	2506B	
眔	KP1−5DCC	7714	25133	
硁	KP1−5FBF	7841	2544B	Non−cognate, should be moved. Ref. L2/25−050
𥞛	KP1−6202	2579B	2E0C4	
𥟙	KP1−622F	257D9	257DF	
簆	KP1−64DA	7C06	25C83	
罀	KP1−68CB	7F80	2E29A	Cognate but NUCV.
荆	KP1−6E6A	8346		 8346 艸 140.6 刀 18.8 G0-3E23 T3-3E57 JMJ-021868 KP1-6E6A 834A 艸 140.6 GE-494C HB1-AFF0 T1-5774 J0-3755 K0-7B2A KP0-F3EC VN-0834A Glyph identical to 834A. Yet 834A occupied by KP0−F3EC.
褁	KP1−73E7	8901	2E590	Cognate but not UCV.
讂	KP1−77D3	8B82	27B66	
𧯕	KP1−7818	27BD5	27BD6	
䥀	KP1−8130	4940	28A6A	 4940 金 167.10 G3-7674 T4-5E2C JMJ-005182 KP1-8130 28A6A 金 167.9 GHZ-64231.12 Glyph different to both. Sim personally thinks 28A6A would be closer.

嶲	KP1−83F7	96DF	5DC2	*reported by Judith*
頚	KP1−8699	981A	2948D	
尳	KP1−8AA9	29A14	5C33	
𩬻	KP1−8B6B	29B3B	2EAC4	
䳹	KP1−8F3D	4CF9	2A13A	

Also, there are some glyph issues in the KP−source glyphs. We do not have the right to revise any of the glyphs since we are not delegates from DPRK; however, among them, there are some identical shapes on different codepoints, which may need to be recorded in UTN #50. One of them was reported by Eiso Chan as mentioned in the first paragraph of this document.

According to the proposals from DPRK to IRG in early years, they use CheongPong V3.0 for the KP0−source characters. Actually, CheongPong V3.0 could be downloaded everywhere on the internet, we can easily see what the glyphs look like in that font, but the main problem is that DPRK did not authorize Unicode to use that font for KP−source representative glyphs, since all the KP0−source characters are included in KPS 10721 (aka KP1), so the KP1−font is used in CodeCharts instead currently.

KP1−3A3A and KP1−5565

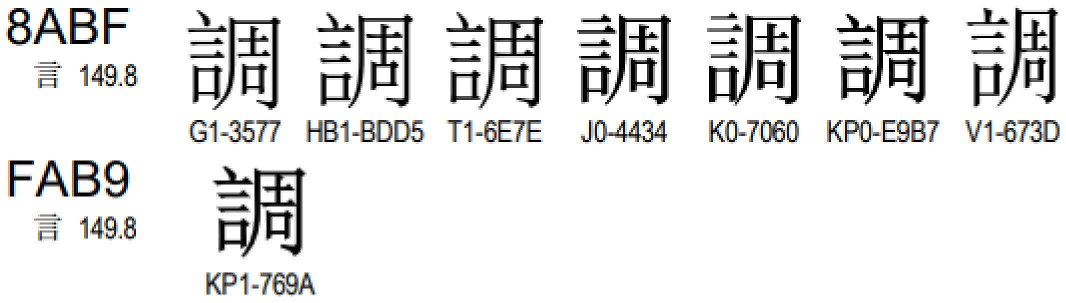


The issue of this character is that, semantically, its radical could be both 口 and 火 − according to *Kangxi Dictionary*, it is both a variant of 呴/欨 (meaning “blow”) and a variant of 煦 (meaning “warm”). *Kangxi Dictionary* listed it under the radical 口, but according to the glyph itself, people may probably think that its radical should rather be 火. This may be the main reason why U+55A3 and U+24275 are currently (Unicode 16.0) under different radicals even they are cognate. Also, KP1−3A3A is under the radical 口 while KP1−5565 is under the radical 火 according to KPS 10721. From the viewpoint of KPS

10721, it may just be a duplicate character caused by negligence. Since we could do nothing on the glyph, and meanwhile 火 & 灬 are even NUCV, I would like to recommend the CJK & Unihan WG to pick one of these strategies:

1. Do nothing on source references, only to record it in UTN #50.
2. Remove KP1−5565 from U+24275, and add a second radical 火 to U+55A3.

KP1−7699 (=KP0−E9B7) and KP1−769A

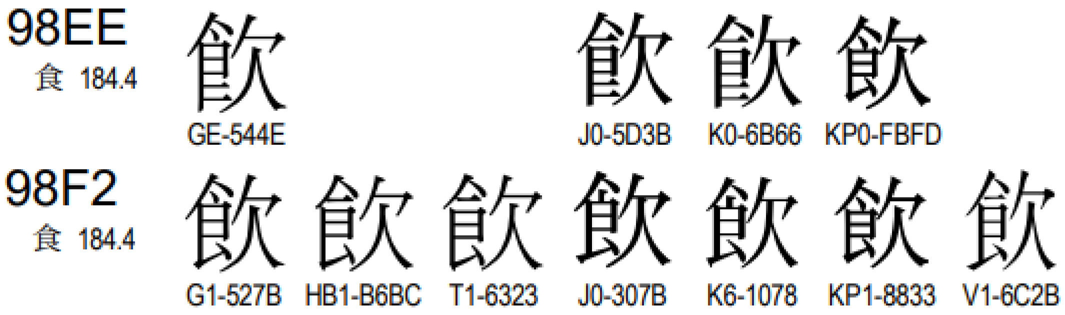


Since this is an issue related to KP0−source, let us just see what it looks like in CheongPong V3.0 (the KP0−font). Also, we could see the other characters with the component 周 in the KP0−font (and the KP1−font is also listed for contrast):

倜凋周彫稠綢調週雕
倜凋周彫稠綢調週雕

It is clear that all the 周s look like ⿵冂⿱丰口 in the KP0−font while all the 周s look like ⿵冂⿱土口 in the KP1−font. Thus it is a systematic issue tentatively could not be solved unless one day we get the permission of using the KP0−font for KP0−source characters in CodeCharts.

KP1−8833 and KP1−8834 (=KP0−FBFD)

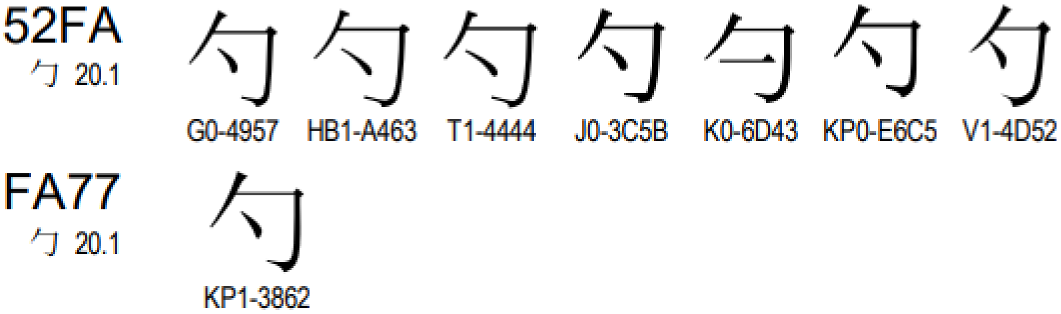


This is also a similar issue as 調 mentioned above. It is a systematic issue for the component 𩙿 . Let us just see some examples:

飢飭飲飯飴飼飽飾餃餉
飢飭飲飯飴飼飽飾餃餉

All the components 𩙿 are designed in an Older Form (something called 旧字形 in Chinese or 旧字体 in Japanese) in the KP0–font, while the same components are designed in a Newer Form (something called 新字形 in Chinese or 新字体 in Japanese). It is really a tricky issue since 飲 and 飮 were encoded separately due to the historical legacy issue. Again, nothing could be done unless one day we get the permission of using the KP0–font for KP0–source characters in CodeCharts.

KP1–3861 (=KP0–E6C5) and KP1–3862


52FA
勺 20.1
G0-4957
HB1-A463
T1-4444
J0-3C5B
K0-6D43
KP0-E6C5
V1-4D52
FA77
勺 20.1
KP1-3862

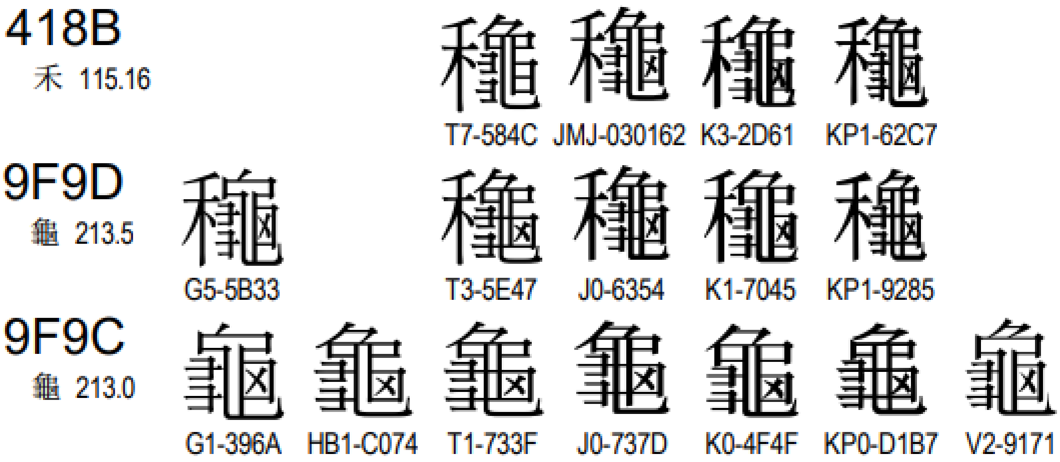
Again an issue similar to 調, but more complicated since the 勺 components are inconsistent within the KP0–font:

勺杓灼的約芍葯豹酌釣
勺杓灼的約芍葯豹酌釣

The KP1–font designs the characters with the component 勺 more consistent, which leads to an issue that the glyph for U+52FA becomes identical to U+FA77. Actually, the orginal KP1–font extracted from the Android application named “Okpyon” does have the 勺 shape, which has to be mapped to U+FA77 (i.e. two shapes swapped). The glyph for U+FA77 was not updated because we wanted to make the KP–source glyphs from the early versions of the Unicode Standard stable – the same issue also occurs for U+FA96 and U+FA99, just as mentioned in the 1st subsection of the 勘误 section in the reconstructed KPS 10721 document. Since we do have the 勺 shape in the KP1–font (although not used in CodeCharts currently), we do not need to “revise” the glyph, I would like to recommend the CJK & Unihan WG to pick one of these strategies:

1. Do nothing on glyphs, only to record it in UTN #50.
2. Change the representative glyph of either U+52FA or U+FA77 to 勻.

KP1–62C7 and KP1–9285

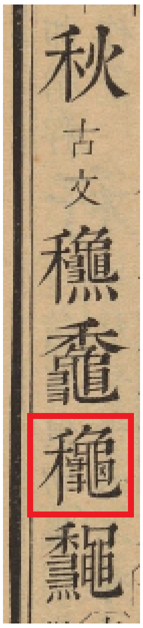

418B
禾 115.16
T7-584C
JMJ-030162
K3-2D61
KP1-62C7
9F9D
龜 213.5
G5-5B33
T3-5E47
J0-6354
K1-7045
KP1-9285
9F9C
龜 213.0
G1-396A
HB1-C074
T1-733F
J0-737D
K0-4F4F
KP0-D1B7
V2-9171

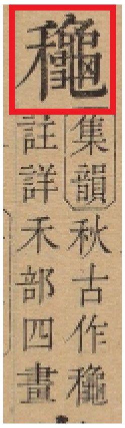
This is the last one, and the most complicated one. It seems that the J–, K– and KP–source glyphs on U+418B should be on U+9F9D (the glyphs for U+9F9C are listed for reference).

For J–source, after checking the update record of Moji Jōhō (aka JMJ), JMJ–030162 was initially on U+9F9D. There were many different glyphs on U+9F9D in JMJ, JMJ–030162 is later moved to U+418B, and all the remaining ones were later registered in IVD. Thus we may just say that, JMJ noticed the unused codepoint (U+418B) and recycled it for one of the variants on U+9F9D to share some pressure. Since it is intended, meanwhile unifiable, we may just ignore its existence.

For K–source, the two glyphs on the two codepoints are even EXACTLY the same. It is strange but I do not know how to explain, maybe we have to ask the Korean experts.

Let us see the most important one, also the main subject of this document, the KP–source. Let us check the *Kangxi Dictionary* first:




龝
集韻秋古作龝
註詳禾部四畫

Some minor differences of the two glyphs could be seen. The left one is under the radical 禾 in *Kangxi Dictionary*, while the right one is under the radical 龜. As expected, U+418B is under the radical 禾 while U+9F9D is under the radical 龜, and the glyph differences of KP1−62C7 and KP1−9285 exactly matched the differences from *Kangxi Dictionary*. Actually there are so many similar cases from *Kangxi Dictionary* but they are all encoded as compatibility ideographs except for this one. Personally I do not have a strong opinion on how to deal with this pair (since we cannot move anything to a new codepoint in the compatibility block), if any from other people, please put forward, otherwise simply record it in UTN #50 would be enough.

(End of document)