Date 2024−12−30

Source:	CheonHyeong Sim (沈天珩)
Title:	Proposal to Disunify U+6138 (憼 & 憼)
Status:	Individual Contribution on IRG #64
Action:	To be considered by China NB, ~~DPRK~~ and IRG

Background

U+6138 looks like the following picture in the CodeCharts of Unicode 16.0:

Fig.1 U+6138 in Unicode 16.0

Obviously, U+6138 mixed the two shapes 憼 and 憼. However, the relationship between these two shapes are the traditional form and the simplified form, thus need to be disunified. In IRGN2608, Eiso Chan once raised this issue but finally no action was taken by IRG.

Rationales of Disunifying

In IRGN2608, Eiso Chan said that,

> KP1−439A is under U+6138 currently, but 来 and 來 are the simplified/traditional pair, and they cannot be unified anytime.

and gave out some evidences of 憼 to prove its existence. Due to space constraints, the evidences in IRGN2608 will not be repeated here. In this proposal, I would like to give out some additional rationales to explain the necessity of disunifying.

First of all, I would like to say that, since DPRK has been inactive in either IRG or WG2 for decades, people in UTC already agreed to make small revisions for the KP−source and record them in UTN #50 without the agreement from DPRK, if the proposed revisions are reasonable. For example, KP1−5B5D was moved from U+24D6A to U+3F94

since Unicode 15.1. Thus we do not need to worry whether we have the right to move KP1−439A away from U+6138.

There are indeed some glyph issues in the KP−source font, but for KP1−439A, it should NOT be a glyph issue but a mapping issue. We could easily get this by the sort of the KPS 10721 character set:

437_惱愖愁愨愨想惰愃惺愯愇愁愃惷慎慅
438_慈愄愰愁愁愁愁愄愁愁愁愁愁愁愁愁
439_愁惠惶愰愁愚愁焦愈愈愁愛意愁惡愕
43A_惹愁傲愁愈愁愁愁愁愁愁愁愁愁愁愁
43B_愁惜愷愁愧愁愁愁愁愁愁愁愁愁愁慎

<div align="center">Fig.2 <u>KPS_10721_reconstructed_by_CheonHyeong_Sim</u></div>

All the characters in Fig.2 are under radical 心. We may try to count the residual strokes, and could easily conclude that, in Fig.2, all the characters before KP1−43A9 are 9, while all the characters after KP1−43AA are 10. Thus the glyph indeed needs to be 来 but not 來.

Also, both 愁 and 愁 appears in the character sets in China:

<div align="center">Fig.3 GB 7590−87（aka G4−source, or 第四辅助集 in Chinese）</div>
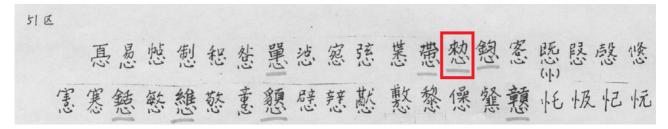
<div align="center">Fig.4 GB/T 13132（aka G5−source, or 第五辅助集 in Chinese）</div>

Note that, the one from Fig.4 (i.e. G5−532D) is exactly the G−source reference on U+6138. It seems to be more reasonble to disunify 愁 and 愁 since they are really treated

as a traditional/simplified pair in Chinese Standards. Thus China NB may consider a horizontal extension for the disunified KP−source character as G4−532D.

Summary

To briefly sum up, we can list 憨 and 憨 in multiple sources as the following table: (Note that, red indicates the current sources for U+6138)

	G−source	T−source	J−source	K−source	KP−source
憨	G5−532D	T3−466E	JMJ−011779	K2−3328	/
憨	G4−532D	/	/	/	KP1−439A

By the Way...

In the feedback to IRGN2608 from TCA, they pointed out that U+208B7 also mixed 来 and 來. However, this was an issue I have alreay raised in IRGN2537, the issue was that, the traditional form never existed in any documentary evidences − just check C00787 in 教育部異體字字典 − it says that, 文獻上多作「齧」, and all the evidences listed on the right side of the page show the simplified from. Thus we would better regard U+208B7 as a glyph issue than a disunify issue. I recommend TCA to revise the glyph.

(End of document)