

Title: **IRGN2422 Request to not disunify U+2F8F0 from U+238A7 and remove U+4DBE from WG2 N5100R3**
 Author: Henry Chan
 Type: Individual Contribution
 Date: 2019/10/21 6pm
 Distribution: IRG
 Pages: 4

Part 1

In IRGN2338 TCA Feedback, TCA requested to disunify

U+2F8F0 (T6-4A3F **𪛗**) from U+238A7 (T6-3857 **𪛗**):

	T-code/UCS-code	USC font	全字庫 font
3	T6-4A3F/U+2F8F0	2F8F0 <small>欠 76.6</small> 𪛗 <small>T6-4A3F</small> ≡ 238A7 𪛗 ~ 238A7 FE00	
	T6-3857/U+238A7	238A7 <small>欠 76.6</small> 𪛗 𪛗 𪛗 <small>UCS2003 GKX-0567.07 T6-3857</small>	

For the group 3, TCA agree U+2F8F0 and U+238A7 should not be unified , need to be encoded separately.

In IRGN2365 Misc Report, it was confirmed as below:

(2) Separately encode some characters listed in IRGN2337 and IRGN2338, except U+2F878 𪛖 (H-8B73). Details are recorded in Appendix 1: Disunifications in IRGN2337 and IRGN2338.

For U+2F8F0 (**𪛗** T6-4A3F), encoded separately/ need a new code point (IRGN2338 TCA TCAResponseRevised)

In WG2 N5100R3, it T6-4A3F is moved to new codepoint U+4DBE as CJK Unified Ideograph:

4DBE
欠 76.6
𪛗
T6-4A3F

However, according to the UCV Rule#173, T6-4A3F 𠄎 and T6-3857 𠄎 should be unified:

173 · unifiable 𠄎𠄎

𠄎 𠄎 𠄎 𠄎

There is also another unification example of 𠄎 & 𠄎 :
 𠄎 U+2F93A is compatibility character which canonicalizes to 𠄎 U+7610:

c. Differences in contact of strokes

c-1. Differences in horizontal connectivity

173 unifiable	𠄎 𠄎 𠄎 𠄎				
	𠄎𠄎 Compatibility Ideographs 𠄎 U+2F93A = 𠄎 U+7610				
	2F93A 𠄎 7610 𠄎 𠄎 𠄎 𠄎				
	𠄎 U+2F859 = 𠄎 U+214E4				
	2F859 𠄎 214E4 𠄎 𠄎 𠄎				
	𠄎 U+2F8F0 = 𠄎 U+238A7				
	2F8F0 𠄎 238A7 𠄎 𠄎 𠄎				

According to MOE dictionary, U+238A7 (T6-3857 𠄎) and U+2F8F0 (T6-4A3F 𠄎) are variants:

C05639 𠄎 解 說 異體字 回查詢結果 複製引用網址

異體字

𠄎 𠄎 𠄎

According to IRG PnP Annex I item I.3, character pair which is cognate shall not be disunified:

I.3 Discouragement of new disunification request

There is a possibility of “pure true disunification” request. This is almost like the new source code separation request. This kind of request shall not be accepted disregarding the reasoning behind. Key difference between “TO BE DISUNIFIED” and “SHALL NOT BE DISUNIFIED” is as follows.

- a. If character pair is non-cognate (meanings are different), that pair of characters is TO BE DISUNIFIED.
- b. If a character pair is cognate (means the same but different shape), that pair of characters SHALL NOT BE DISUNIFIED.

Disunification request with reason of mis-application (over-application usually) of unification rule should NOT be accepted due to the principle in resolution [M41.11](#).

Therefore, U+2F8F0 (T6-4A3F 𠵶) should not be disunified from U+238A7 (T6-3857 𠵶).



I suggest that

be removed from WG2 N5100R3.

Part 2a

Because T6-4A3F (𡗗) is the more common preferred form, and T6-3857 (𡗗) is a less common form, I suggest that the two glyph and source references be swapped, i.e.

T6-4A3F (𡗗) be moved to U+238A7, and T6-3857 (𡗗) moved to U+2F8F0.

Part 2b

I also suggest China to consider to update the glyph of U+238A7 to 𡗗, which is more common form, with the source reference revised to GHZR-42297.07:

𡗗同“𡗗”。《玉篇·欠部》：“𡗗”，同“𡗗”。

(Hanyu Dazidian Vol 4 page 2297 entry 7)

Summary of changes from Part 2a & 2b:

	Before Proposed Changes	After Proposed Changes
U+238A7	238A7 欠 76.6 𡗗 𡗗 𡗗 UCS2003 GKX-0567.07 T6-3857	238A7 欠 76.6 𡗗 𡗗 𡗗 UCS2003 GHZR-42297.07 T6-4A3F
U+2F8F0	2F8F0 欠 76.6 𡗗 T6-4A3F	2F8F0 欠 76.6 𡗗 T6-3857