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

		T-code/UCS-code	USC font	全字庫 font
·	3	T6-4A3F/U+2F8F0	2F8F0 欠 76.6 ——————————————————————————————————	欺 欸 欸
		T6-3857/U+238A7	238A7 ½ 76.6	東大

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 屮 (H-8B73). Details are recorded in Appendix 1: Disunifications in IRGN2337 and IRGN2338.

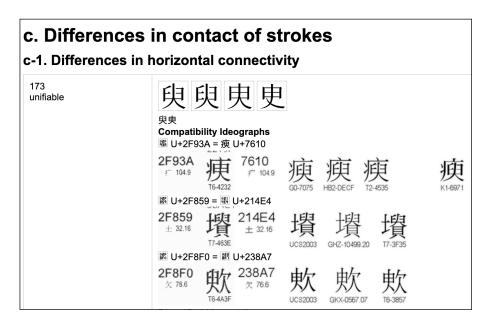
For U+2F8F0 (T6-4A3F), encoded separately/ need a new code point (IRGN2338 TCA TCAResponseRevised)

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



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





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



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

1.3 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) should not be disunified from U+238A7 (T6-3857).



I suggest that

be removed from WG2 N5100R3.

Part 2a

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

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

Part 2b

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

「「大」。《玉篇·欠部》:"敟",同"炊"。

(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	238A7 欠 76.6	
U+2F8F0	2F8F0 欠 76.6 T6-4A3F	2F8F0 欠 76.6 下 76.3857	