Source:	China
Author:	Tao Yang
Title:	Feedback on IRG N2840
Meeting:	IRG #65
Actions required:	To be considered by IRG
Distribution:	IRG
Medium:	Electronic
Page:	4
Appendix:	None

Thank Mr. Ma Shijie for pointing out a quite important issue that, the Duplicate-encoded ideographs and erroneous extended glyphs resulted in a small amount of glyph duplication, which would bring more challenges to NLP.

According to the original sources, China considered the suggestion from Mr. Ma and response as below:

1. The glyph revision of G-source in URO should be considered carefully, the glyph matches the evidence should be kept in the code point.

2. The situation where two code points using the same glyph should be corrected as much as possible. Mr. Ma did a very valuable job by listing a series of duplicated characters.

3. Analysis for glyph revision：
3.1 U+34A8 僨/U+20457僨

Mr. Ma suggests to change the glyph of U+34A8 to and change the references both.

U+34A8 is a character in the double byte area of the Chinese coded character set, which is included in GBK in China's information technology standards. GBK has been used in China for 30 years, and its application is very stable and extensive. Characters other than GBK are not widely supported by various information systems. Therefore, in general, I hope not to adjust the font in this range and maintain the basic stability of characters.

The glyph of U+34A8 僨 is with the reference G5-3329, which shows the original glyph

 is almost the same with the encoded one. The lower-left part was corrected from to which derived from the small seal component .

The answer is not to change the glyph and reference of U+34A8 僨, but change the glyph of U+20457 to with the reference changing from GKX-0121.39 to GHZR-10281.02.

3.2 U+821D 舝/U+246C9 舝

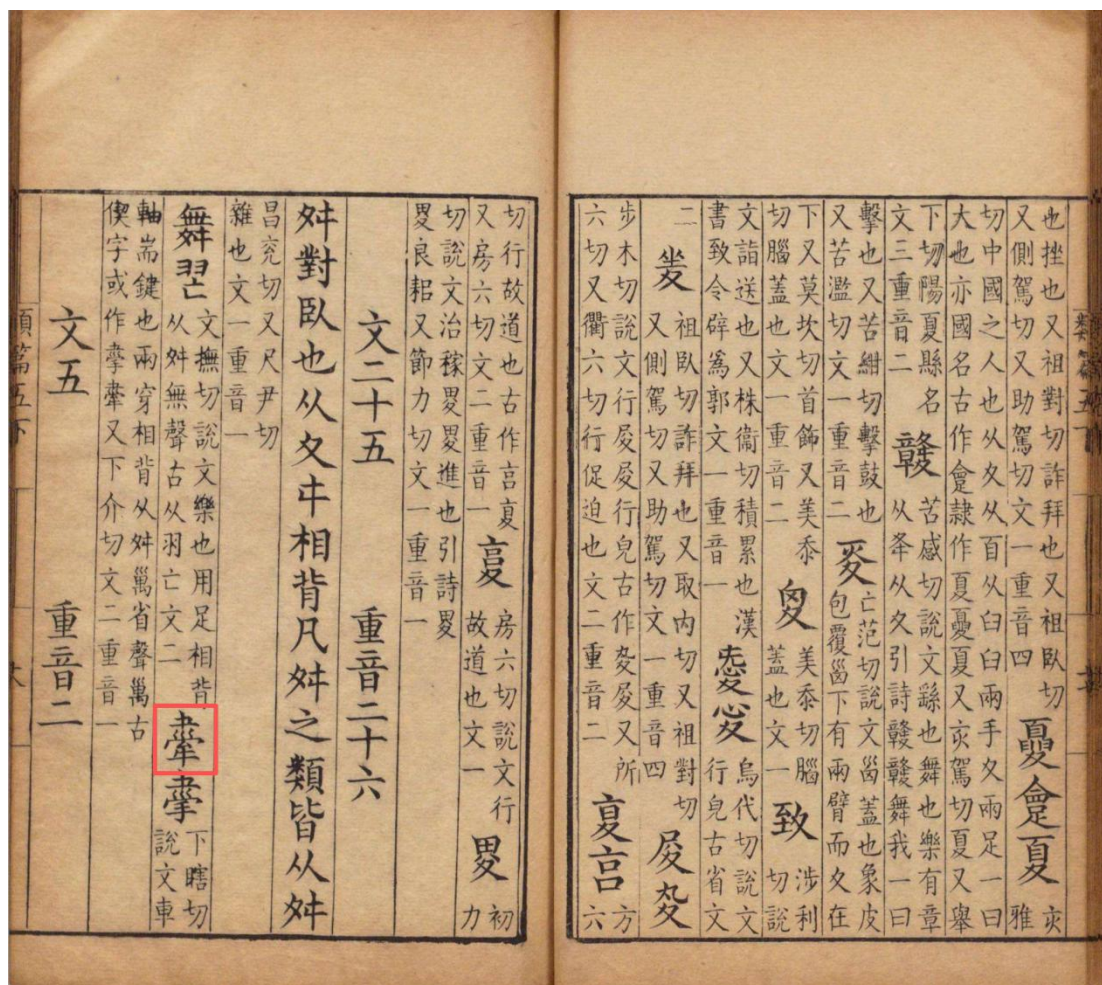Mr. Ma suggests to change the glyph of U+246C9 to and change the references both.
The glyph of U+821D 舝 matches the original evidence and G5 is an authority source for UCS, so that the reference of U+821D 舝 doesn't need to be changed.

As the glyph of U+821D 舝 has been determined to be , so U+246C9 should be encoded as the variant , and the reference has been changed to GCESI-00535 in IRGN2864.

3.3 U+204F2夵/U+23515㔕

Mr. Ma suggests to change the glyph of U+204F2 to 㔕 and change the reference to GHZR-10237.05, and change the reference of U+23515 to GHZR-10237.06.

China accept that suggestion.

In the meanwhile, China noticed that the radical of U+204F2 should be the same with U+23515.

So China wants to change the radical of U+204F2 from 入 11.10 to 木 75.8.

3.4 U+249BC䦼/U+249E9䦼

Mr. Ma suggests to change the glyph of U+249BC from 䦼 to 䦼, change the reference to GHZR-21200.09.

China accept that suggestion.

As the glyph of U+249BC has been determined to be 䦼, so U+249E9 should be encoded as the variant 䦼, and the reference has been changed to GHZ-21130.04 in IRGN2864.

Code point	UCS	New evidences	Suggested glyph	Suggested reference	Revised glyph	New references
U+34A8	僲 僲 G5-3329 T7-574C	僲同"仙"。《集韻·僲韻》："僲,隷作僊,通作仙。" 《漢書·郊祀志》:"(求)僲人羨門之屬。"顔師古注:"古亦以僲為仙字。"	僲	GHZR-10281.02	No change	No change
U+20457	僊 僊 僊 GKX-0121.39 T4-662F JMJ-056959	僊膚 説文·人部 同"仙"。《廣韻·仙韻》："僊","仙"的古文。	--	GHZR-10280.06	僊	GHZR-10281.02
U+821D	羍 羍 羍 G5-3039 HB2-DFA3 T2-4648	羍 [xiá 《广韵》胡瞎切，入鎋，匣。] ❶车辖。古代为固定车轮而插在车轴两端的键。《诗·邶风·泉水》:"載脂載羍,還車言邁。"孔颖达疏:"古者車不駕則脱其羍……今將行,既脂其車,又設其羍。"又《小雅·車羍序》:"車羍,大夫刺幽王也。"陆德明释文:"車軸頭鐵也。"❷星名。《史记·天官书》:"房爲府,曰天駟。其陰,右驂。旁有两星曰衿;北一星曰羍。"	--	GHC-010797.02	No change	No change
U+246C9	羍 羍 羍 UCS2003 GHC T5-4730	羍 xiá 音狹❷ 同"辖",穿在车轴两端孔内使车轮不脱落的键。字见《篇海》。	羍	GHZ-0854.55	羍	GCESI-00535
U+204F2	桀 GHZ-10199.02	桀 同"乗"。《篇海類編·人事類·人部》:"桀,通作乗。"	桀	GHZR-10237.05	桀	GHZR-10237.05
U+23515	桀 桀 桀 GHZ-21235.06 T6-4A39 JMJ-038092	桀 桀 説文·桀部 同"乗"。《龍龕手鑑·入部》:"桀,古文。車桀也。"	--	GHZR-10237.06	--	GHZR-10237.06
U+249BC	瑿 瑿 T5-473E JMJ-041022	瑿 瑿 説文·玉部 《说文》:"瑿,石之似玉者。从玉,臤聲。" wàn 《集韻》烏貫切,去换影。元部。 似玉的石。《説文·玉部》:"瑿,石之似玉者。"	--	--	HE	GHZ-21130.04
U+249E9	瑿 GHZ-21130.04	瑿 瑿 説文·玉部 《说文》:"瑿,石之似玉者。从玉,臤(旧字形作'臤')聲。" wàn 《集韻》烏貫切,去换影。元部。 似玉的石。《説文·玉部》:"瑿,石之似玉者。"	瑿	GHZR-21200.09	瑿	GHZR-21200.09

(EOF)