IRGN2179 (IRG WS 2015 Version 3.0) Review	
Title	Review Feedback on Consolidated Comments & Supplementary Information to Comments
Author	Henry Chan
Date	2017/06/01
Type	Individual Contribution to IRG

Issue 1

<u>02179 UTC-02651</u>

Henry's Comment
02179 UTC-02651 = 烱 (U+241C6) / 烟 (因 ~ 曰 -- 第一批异体字整理表)

Japan's Comment
Unify with U+241C6 烱

UK's Comment
Disagree. We do not believe that 回 and 曰 are unifiable components.

Henry's Comment
It is common for the middle of 回 to be written as 曰 in print, such as:

although the reverse is rather uncommon.

In the evidence provided, it is given that 02179 is a variant of 烟:

Fig. 1393. *Zhonghua Zihai* (Beijing, 2000) p. 953 col. B

To also quote from MOE Dictionary
(http://dict2.variants.moe.edu.tw/variants/rbt/word_attribute.rbt?quote_code=QTAyNDIwLTAwMQ):

炟 is a variant of 烟 and that the pair 因 / 曰 is included in 《第一批异体字整理表》. Therefore, the equivalence relationship between 02179 and 炟 is beyond reasonable doubt.

Suggested Action Item
Unify/IVD

URL: https://github.com/hfhchan/irg-ws2015/issues/1

Issue 2

Henry's Comment
03798 UTC-01941 UNIFY to 禅 (U+891D)

UTC, HK Comment
Unify with U+891D 禅

UK's Comment
Disagree. We think that the unification of the components 单 单 for U+7985 禅 was a mistake, and causes problems for users when a default font shows an unacceptable glyph form. The G-source glyph for U+891D has 单 on the right, so font developers will follow this glyph form when designing fonts for PRC, but the G-source glyph form for U+891D is unacceptable as the simplified form of U+891D 襌.

Therefore we strongly think it will serve users best to encode UTC-01941 (⯾礻 单) as a separate character.

Henry's Comment
The above problem will occur for U+20219 俴 and U+2548E 碊 also, even though they occur in Extension B. However, China has agreed to correct multiple erroneous glyphs in GE standard in IRGN 2170 (involving U+8669 and U+3B9D). Therefore, the rejection to correct U+891D in line with China's normalized glyphs should not be accepted.

Action Item
Unify

URL: https://github.com/hfhchan/irg-ws2015/issues/2

Issue 3

00597 UTC-02810

00597	□ 30.8		嗲 UTC-02810 (UK)
	□ 口 参		
折	N/A		

UTC's Comment
Unify with U+2D227 □

2D227
□ 30.8
嗲
GZ-0581203

UK's Comment
Agree.

Henry's Comment
Disagree. 叁 and 参 are cognate but have different semantics in modern day. The phonetic of U+2D227 should be confirmed to be 参 instead of 叁 before unification.

Action Item
Disunify or Postpone.

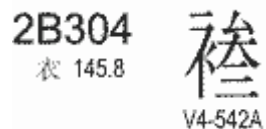URL: https://github.com/hfhchan/irg-ws2015/issues/3

Issue 4

03800 UTC-01942

03800	衣 145.8					褖	
	𪐷衤参					UTC-01942	
	折	N/A				(UK)	

UTC, HK Comments

Unify with U+2B304 褖

2B304
衣 145.8
V4-542A

UK Comment

Agree.

Henry's Comment

Disagree.

The pronunciation of U+2B304 褖 is given to be tươm, 叁 to be tam, and 参 to be tham/sam/sâm/khươm on the Nom Foundation Nom Lookup Tool. It is hightly probable that the phonetic of U+2B304 褖 is 叁 instead of 参.

Before the phonetic of U+2B304 褖 can be truly confirmed, U+2B304 and 03800 should not be unified.

Action Item

Postpone or Disunify.

URL: https://github.com/hfhchan/irg-ws2015/issues/4

Issue 5

05541 UTC-02614

Henry's Comment

IVD to U+2B733 (冉 ~ 冄)? PRC Conventions prefer 冉 > 冄.

UK's Comment

Disagree. We do not think that 冄 and 冉 are unifiable components, and 《漢語大字典》 has separate entries for both U+4DB2 䶲 and U+2A6AE 𪚮 and the two corresponding simplified forms.

Henry's Comment

The unification of 冄 and 冉 should be considered unifiable for IVD. There are already many exact equivalents encoded in URO, and many more can be encoded.

Examples of exact equivalents:
U+5189 冉 = U+5184 冄
U+67DF 柟 = U+678F 枏
U+8043 聃 = U+803C 耼
U+86BA 蚺 = U+86A6 蚦
U+88A1 袡 = U+887B 衻
U+9AEF 髯 = U+9AE5 髥
U+59CC 娌 = U+36A9 姩
U+8211 舑 = U+4459 舙
U+82D2 苒 = U+44A3 苳
U+279A6 訵 = U+46C1 詽
U+294FF 顃 = U+4AC7 頶
U+5465 呥 = U+20BCD 呍
[...]
U+722F 爯 = U+2DDAA ☐
U+7A31 稱 = U+2E0CE ☐

Action Item

Unify via IVD.

URL: https://github.com/hfhchan/irg-ws2015/issues/5

Issue 6

00069 UTC-02765

Henry's Comment
= 几 (U+20627) (variant, protrousion of strokes.)

UK's Comment
Disagree. Non-cognate, and stroke variation is significant.

Henry's Comments
Below is UK's evidence of 00069:

Fig. 1514. *Hanyu Fangyan Da Cidian* (Beijing, 1999) p. 143

几 shū ㄕㄨ

【几几】❶后缀。用于形容词后，表示程度。兰银官话。甘肃武威。李鼎超《陇右方言·释言》："今谓短曰'短～'、肥曰'肥～'，音近'出'。" ❷〈形〉形容动物（多指短羽鸟禽）飞行快速且常作停歇的样子。㊀冀鲁官话。河北沧县。1933年《沧县志》："～，此状辞也。凡物之行走速而屡作停顿，其貌均谓之～。" ㊁西南官话。四川荣县。1929年《荣县志》："短羽飞曰～。"

Below is HYDZD evidence of 几 (U+20627):

几乚。¨

几 𠘧 説文・几部

《説文》："几,鳥之短羽飛几几也。象形。"
shū 《廣韻》市朱切,平虞禪。侯部。
❶〔几几〕1.短羽鸟飞貌。《説文・几部》："几,鳥之短羽
飛几几也。"饒炯部首訂："几几者,形容短羽鳥飛也。"
2.病貌。《字彙補・几部》："几,病貌。"《素問・刺腰痛
篇》："腰痛俠脊而痛,至頭几几然,目晥晥欲僵仆。"《傷
寒論・辨太陽病脈證並治法上》："太陽病,項背強几几,
反汗出惡風者,桂枝加葛根湯主之。"成無己注:"几几,
伸頸之貌也。"

First, the pronunciations are the same.
Second, the second meaning of 00069 is the same as the strict meaning of U+20627.
Third, the first meaning of 00069 is used as a grammatical suffix to show extent. Very likely, it is a character borrowed for its sound.
Lastly, the top left part shape of 00069 matches the Shuowen shape of U+20627. 00069 is likely simply another transcription of the same character.

It is likely they are the same character.

Action Item
Unify or Postpone.

URL: https://github.com/hfhchan/irg-ws2015/issues/6

Issue 7

Character Stroke Count for 丽

丽 is present in the Kangxi Dictionary with a SC = 7. The radical is Dot (丶). Therefore, the total SC should be 8.

It is also counted as total SC = 8 when on top of 鹿

However, U+4E3D in the Code Charts has an SC = 6 (total SC = 7)

4E3D
— 1.6
G0-4076   H-8946   T3-2740

Affected Characters as follows:

- 02746 UTC-01877

02746	目 109.7		瞓 UTC-01877 (UK)
	回目丽		
	橫	N/A	

- 04094 UTC-02120

04094	足 157.7		蹦 UTC-02120 (UK)
	回足丽		
	橫	N/A	

- 00376 UTC-01690

00376	刀 18.7	刪
	〼 丽 刂	UTC-01690
	橫	N/A	(UK)

- 03824 UTC-02112

03824	見' 147'.7	覼
	〼 丽 见	UTC-02112
	橫	N/A	(UK)

- 03793 UTC-01940

03793	衣 145.7	禰
	〼 衤 丽	UTC-01940
	橫	N/A	(UK)

- 01398 UTC-01809

01398	手 64.7	搗
	〼 扌 丽	UTC-01809
	橫	N/A	(UK)

Action Item

Add 丽 to IRGN954AR with total SC of 8.

URL: https://github.com/hfhchan/irg-ws2015/issues/7

Issue 8

Unification of 幸, 幸 and 羍

In UCS, there are numerous examples of 幸, 幸 and 羍 disunified.

There are three different etymologies for characters that contain 幸.

(1) U+3694 幸 (niè, handcuffs).
Examples of characters include (usually as a semantic component):

- U+57F7 執 / U+2163A 執 / U+21655 靮 / U+2065C 靮 / U+26383 埶
- U+5831 報 / U+21648 報
- U+776A 睪 / U+251E1 睪
- U+25216 睟
- U+23582 椌
- U+20DBF 嗌
- U+2676F 膝
- U+260A1 緈
- U+26051 緈
  (etc)

(2) U+21D18 㐷 (xìng, fortune)
Examples of characters include (usually as a phonetic component):

- U+200B7 㣷 (alternative transcription) of 㐷
- U+5548 哖 / U+20D43 哖
- U+6DAC 涬 / U+23DDF 涬
- U+46ED 諲 / U+27A2B 諲
- U+7DC8 緈 / U+2609C 緈 / U+260C9 緈
  (etc)

(3) U+7F8D 羍 (dá, small sheep)
Examples of characters include (usually as a phonetic component):

- U+5548 哖 / U+20D43 哖
- U+9054 達 / U+9039 達

From the above examples, it can be shown that the shape difference between 幸, 幸 and 羍 is not generally representative of a systematic semantic difference in Kaishu. In the case of 哖/哖/哖, the

dictionary meaning/pronunciation of the characters is actually opposite to the normative meaning of its phonetic/semantic symbol.

Actually, these three forms were never really distinguished in handwriting. It would be distinguished by context. There is no need for multiple variants of the same character to be encoded. Trying to map every single variant into a dictionary into UCS would only cause confusion to the exact semantic meaning. IVD can be used for the preservation of exact shape.

The following characters are the semantically equivalent to corresponding encoded ideographs, and thus should be unified in WS2015:

- 00682 UTC-01451
  Semantic Origin: 幸 - handcuffs
  UNIFY TO U+5709 圉

- 00195 USAT09927
  Semantic Origin: 㚈 (phonetic)
  UNIFY TO U+5016 倖

- 00882 USAT09928
  Semantic Origin: 㚈 (phonetic)
  UNIFY TO U+5A5E 婞

- 03375 GHZR52968.20
  Semantic Origin: 幸 (semantic of phonetic)
  UNIFY TO U+26525 𦔥

- 03469 GHZR42270.04
  Semantic Origin: 幸 (semantic of phonetic)
  UNIFY TO U+443E 䐾

- 02677 T13-2E70
  Semantic Origin: 幸 (semantic)
  UNIFY TO U+24FF9 𤿹

URL: https://github.com/hfhchan/irg-ws2015/issues/8

Issue 9

01416 UTC-02632

WS2015 v.3 Discussion Record
unified by U+2BF4A (GZFY-00688) for IVS, irg47.

01416	手 64.8		撑 UTC-02632 (UK)	IRGN2179Unified&WithdrawnV3.0 unified by U+2BF4A (GZFY-00688) for IVS, irg47.
	回扌 回䒑子			
	竪	N/A		

Henry's Comment
unified with U+632C for IVS, irg47. NOT unified with U+2BF4A (GZFY-00688).

UK Response
Disagree. It is a non-unifiable variant of U+632C, and should be encoded separately. As reported in IRGN2108Andrew_WG2N4682.pdf, the glyph form of U+2BF4A is incorrect, and should be corrected to 口口

扌学, so UTC-02632 cannot be unified with U+2BF4A.

Henry's Additional Comment
According to my meeting notes, the resolution was unification with U+632C for IVS instead of U+2BF4A (GZFY-00688).

--
Additional Info:

Fig. 1373. *Zhonghua Zihai* (Beijing, 2000) p. 344 col. B

撑 同"挗"。见《龙龛》。

1

632C 手 64.7 挗 挗 挗 挗 挗
G5-4C30 HB2-D1B4 T2-2F36 K2-352D V0-3764

URL: https://github.com/hfhchan/irg-ws2015/issues/9

Issue 10

WS2015v3.0 Discussion Record
IRGN2179PostponedV3.0
pending for solutions (not unified by U+82B2, G source of U+82B2 may be changed), irg47.
unified by U+82B2, irg46.

Henry's Comment

- U+82B2 is used in 第一批异体字整理表 and 《通用規範漢字表》異體字 to mean 花; Keep G-source of U+82B2 unchanged.
- 03555 should be separately encoded.

UK Response
Strongly agree!

Action Item
Suggested to annotate the Code Charts with the correct semantics

82B2
艸 140.4
GE-3A38   T3-2A6C   J14-7621

● Variant of 花 U+82B1

XXXXX
艸 140.4
UTC-01950

● Simplification of 薔 U+83D5

URL: https://github.com/hfhchan/irg-ws2015/issues/10

Issue 11

[removed]

Issue 12

00698	土 32.5					坂 UTC-01204	
	□土夌						
	撇	N/A					

Henry's Comment:

00698 UTC-01204 UNIFY TO 坂

UTC's Comment:

IDS is □土夌. The UTC does not agree.

Henry's Supplementary Information:

夋 and 夌 are variant forms of the same component*. In another version of 弇山堂別集卷三十三, the character 坂 (U+577A) is used instead:
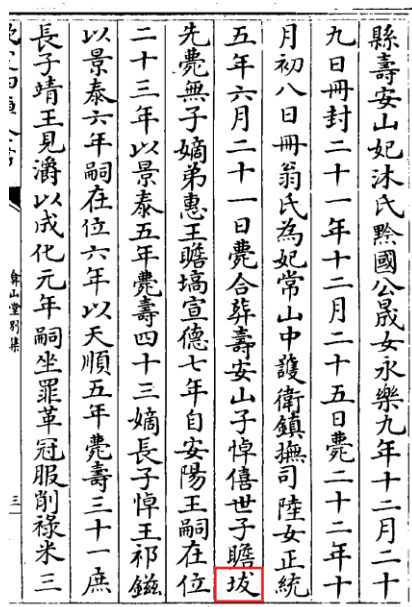
Source: http://ctext.org/library.pl?if=gb&file=58118&page=103, which is 《欽定四庫全書》本·史部五·雜史類。

*: Referencing variants of 拔 (U+62D4) from the MOE Dictionary:

The source quoted by MOE Dictionary has been omitted for brevity.

A list of characters containing 犮 are listed as follows:

U+53D0　　叐: Kanxi / Hanyu Dazidian: variant of 犮

U+39DE　　� 抜: variant of 拔based on evidence of 拔 in MOE Dictionary

U+47E6　　跋: variant of 跋 based on evidence of 跋 in MOE Dictionary

U+2209B　　㈊: unverifiable; G4K-sourced character; original sources could not be found

U+2342A　　枝: TF-sourced character; pronunciation same as 枚

U+24923　　玐: TF-sourced character; pronunciation same as 玫

 U+2595C　　窡: variant of 突 according to Hanyu Dazidian

U+25FC8　　綏: variant of 綏 @MOE Dictionary, HYD and Koseki

U+26B5E　　茇: variant of 茇 @ HYD

U+296BF　　飰: variant of 飰 @ MOE Dictionary, HYD and Koseki

U+2989A　　駁: variant of 駁 @ KX, MOE, HYD, Koseki

U+2C4B2 袯: TC-sourced character; pronunciation same as 被

U+2CAC6 鈸: TD-sourced character; pronunciation same as鈸

U+2CF74 伐: variant of 伏 according to SAT Database

U+2D71E怭: used in name of person in SAT database or as a phonetic transcription; meaning could not be verified.

U+2D805: variant of 戻 according to Mojikiban

2D805
戸 63.5
JMJ-057548　USAT-05952

U+2DF7D: variant of 鉢 according to SAT Database, which is variant of 盋 according to PRC 第一批异体字整理表, Koseki & HYD

2DF7D
皿 108.5
USAT-00269

U+2E0BA: variant of 秡 according to Mojikiban "読み・字形による類推" with remark "地名外字"

2E0BA
禾 115.5
JMJ-059983

U+2E2DF: variant of 黻 according to footnote in SAT Database.

2E2DF
而 126.8
USAT-01944

U+2E3DF: variant of 泼 according to footnote in SAT Database. The next character was 若, so a possible phenomenon of a "類化"

2E3DF	
艸 140.8	菝
	USAT-03904

To summarize, in majority of cases, 犮 component is a strict variant of 犮. In 2 cases it is equivalent to 犬, which is a property also shared by 犮 itself. In 2 cases the source of this character could not be verified, and in 1 case there were multiple sources to show that it was a "corruption" of another similar shaped component.

Therefore I think there is enough evidence to regard 犮 as unifiable to 犮.

In fact, this rule is regarded as a normalization by ROK in IRGN2154:

| 8-3 | 犮 | 犮 | 拔-拔[B094;350c;3(丙)], 跋-跋[B061;278a;3(乙)], 軷-軷[B104;218a;4(丙)], 祓-祓[B112;133d;4(丙)], 魃-魃[B102;271b;6(丙)], 紱-紱[A273;16a;3(丙)], | same above. |

Conclusion:

Based on the above evidence(s), I believe 00698 UTC-01204 should be unified / IVD to 坺 (U+577A).

Affected characters:

00698 UTC-01204 unify/IVD to 坺 U+577A

00861 USAT90292 unify/IVD to 妭 U+59AD

URL: https://github.com/hfhchan/irg-ws2015/issues/12

Issue 13

04035	貝 154.14					贆	
	□貝賓					UTC-01167	
	點						

Henry's Comment:

04035 UTC-01167: more evidence? (賓=賓) no □貝賓 exists yet.

UTC's Comment:

The UTC does not agree. The supplied evidence is clear.

Henry's Additional Comments:

賓 is a strict transcription of the character which is more popularly written as "賓" in modern times. The evidence from Grammata Serica Recensa shows only the transcripted form and not the original Oracle Bone or Bronze or Seal Script:
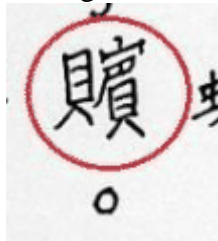
This transcription has caught my attention because no character composed of □貝賓 exists yet – so it is possible that UTC-01167 is an incorrect transcription of a certain character, or the character has completely vanished in medieval and modern usage so no modern transcription exists.

If possible, the original source that this transcribed character was based on should be provided, so the transcription can be verified.

Nevertheless, given the historical significance of Grammata Serica Recensa, unless the original oracle/bronze/seal sources show clear evidence that this character has been horribly mis-transcribed, the character in its current form is still worth encoding.

URL: https://github.com/hfhchan/irg-ws2015/issues/13

Issue 14

00777	土 32.12					塡 UTC-01219	
	▢土▢直木						
	横	N/A					

00765	土 32.11					塡 UTC-01217	
	▢土▢直木						
	横	N/A					

Henry's Comment:

00765 UTC-01217 UNIFY WITH 00777 UTC-01219 (keep 00777).

UTC's Comment:
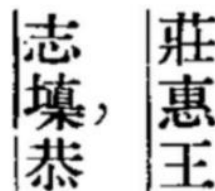
IDSes are ▢土▢盲木 and ▢土▢直木. The UTC does not agree.

Henry's Additional Comments:

The sources provided by UTC are as follows:

Fig. 128:《南明史》第四册, Page 817 [UTC-01217]

Fig. 130:《明史》卷一百, Page 2512 [UTC-01219]

Given the evidence provided by the UTC, it is rather obvious from the context that they should be referring to the same person, and thus be the same character.

A Wikipedia article shows that according to 《明實錄》, the name of the person is 志塡. If this can be confirmed by any expert familiar with the members of the Ming dynasty imperial family, then the two corrupted forms (UTC-01217, UTC-01219) should be unified to 塡 (U+365E) (via IVD).

《明實錄》 can be found on 中央研究所 歷史語言研究所 明實錄、朝鮮王朝實錄、清實錄資料庫:

URL: https://github.com/hfhchan/irg-ws2015/issues/14

Issue 15

Multiple Glyphs were withdrawn by the submitter but were not reflected in the Working Set:

- 00123 UTC-01423: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00130 UTC-01318: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00138 UTC-01391: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00296 UTC-01326: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00345 UTC-01329: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00348 UTC-01330: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00475 UTC-01337: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00524 UTC-01421: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00542 UTC-01338: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00560 UTC-01339: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00561 UTC-01369: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00662 UTC-01342: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00814 UTC-01345: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00827 UTC-01372: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00857 UTC-01353: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00863 UTC-01355: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00866 UTC-01441: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00871 UTC-01354: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00874 UTC-01356: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00898 UTC-01358: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00933 UTC-01437: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00938 UTC-01362: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00953 UTC-01367: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00966 UTC-01368: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00969 UTC-01363: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 00970 UTC-01366: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01152 UTC-01428: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01186 UTC-01349: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01194 UTC-01371: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01319 UTC-01314: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.

- 01320 UTC-01373: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01331 UTC-01374: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01336 UTC-01387: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01493 UTC-01390: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01520 UTC-01377: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01533 UTC-01341: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01703 UTC-01442: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01721 UTC-01379: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 01850 UTC-01384: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02187 UTC-01378: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02193 UTC-01386: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02333 UTC-01388: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02387 UTC-01480: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02446 UTC-01400: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02541 UTC-01392: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02563 UTC-01457: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02641 UTC-01399: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02684 UTC-01393: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02843 UTC-01396: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 02936 UTC-01397: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03026 UTC-01398: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review; SC - 21.
- 03332 UTC-01401: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03349 UTC-01402: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03355 UTC-01403: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03384 UTC-01405: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03487 UTC-01406: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03670 UTC-01404: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03679 UTC-01411: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03774 UTC-01412: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03775 UTC-01413: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03835 UTC-01415: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03969 UTC-01416: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 03981 UTC-01418: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.

- 04685 UTC-01424: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 05061 UTC-01425: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 05094 UTC-01408: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.
- 05222 UTC-01350: Withdrawn by submitter in WS2015v2 - IRGN2155 UK Review.

URL: https://github.com/hfhchan/irg-ws2015/issues/15

Issue 16

There are multiple normalization issues with the Zhuang characters submitted by the Guangxi University. Such as, 橫 should always be changed to 提 on the left side, but they are not in the Zhuang characters. In many cases, the evidence submitted is in the correct normalized form, but the font provided by Guangxi University is not.

Once they are coded, it is very troublesome to change the representative glyph. Therefore, it is suggested that Guangxi University normalize the Zhuang characters properly before their submission.

01594 G_Z3561201: left side 星 does not follow PRC conventions - should be a 提 not 橫

01618 G_Z3551104: left side 星 does not follow PRC conventions - should be a 提 not 橫

04883 G_Z0721301: Does not match PRC conventions. Compare with 養.

02445 G_Z1402302: Does not match PRC conventions. Second stroke of 馬 should not be joined with the 6th stroke.

05239 G_Z0211201: Does not match PRC conventions, last stroke of left component should be 點

03354 G_Z2231201: Does not match PRC conventions, last stroke of left component should be 點

00315 G_Z3842301: Does not match PRC Conventions, last stroke of left component should be 點

04621 G_Z2382304: Does not match PRC Conventions: The third stroke of 犬 should be 點, not 捺.

00065 G_Z1652501: Does not match PRC Conventions: The last stroke of 及 should be 點, not 捺; or the structure should be changed to enclosure.

00264 G_Z4291302: Does not match PRC Conventions: Right hand side should be 彡 (⊟尤彡).

00523 G_Z2042303: Does not match PRC Conventions: last stroke of top left component should be 點.

00629 G_Z2302202: Does not match PRC Conventions: last stroke of top left component should be 點.

00534 G_Z1592101: Does not match PRC Conventions: last stroke of left component should be 點.

00536 G_Z0811201: Does not match PRC Conventions: last stroke of left component should be 提.

00659 G_Z1831401: Does not match PRC Conventions: last stroke of top left component should be 點.

01147 G_PGLG2017 doesn't match PRC conventions, last stroke of left component should be 點.

03527 G_Z2181407 doesn't match PRC conventions; last stroke of left component should be 提.

01149 G_Z3112502 doesn't match PRC conventions, last stroke of left component should be 點.

01150 G_Z0431401 doesn't match PRC conventions, last stroke of left component should be 點.

03665 G_Z1501101 doesn't match PRC conventions, last stroke of left component should be 提.

03805 G_Z1202503 doesn't match PRC conventions, last stroke of left component should be 點.

03961 G_Z2582201 doesn't match PRC conventions, fourth stroke of left component should be 點.
03974 G_Z1412404 doesn't match PRC conventions, last stroke of left component should be 提.
05311 G_PGLG3052 doesn't match PRC conventions, fifth stroke of left component should be 點.
02577 G_Z1432204 doesn't match PRC conventions, last stroke of left component should be 點.
02581 G_Z2782104 doesn't match PRC conventions, last stroke of left component should be 點.
02328 G_Z1602601 doesn't match PRC conventions, last stroke of left component should be 點.
03144 G_Z3651201 doesn't match PRC conventions, last stroke of left component should be 點.

URL: https://github.com/hfhchan/irg-ws2015/issues/16

Item 17

00346 G_Z1841301

G_Z1841301 恐
《古壮字字典》第184页左栏第3条
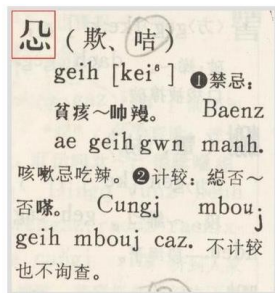
The 几 (U+51E0) is the phonetic, not 几 (U+20627). The character 00346 cannot be changed to use 几 (U+20627). It should use 几 (with hook).

The presence or absence of the hook is not a location variant form issue. Refer to U+28972 where the hook is present even when the component is situated at the top:

URL: https://github.com/hfhchan/irg-ws2015/issues/17

Issue 18

The following submitted Zhuang characters do not use components that are considered standard by PRC. PRC may wish to say that Zhuang characters are not expected to be normalized. However, Ideographs used by Han languages and dialects may use the same character, if such evidence is discovered, then those characters will likely be unified (or IVD) with the Zhuang characters. Then, the existing Zhuang glyghs will need to be modified to use the PRC normalized form, creating lots of problems for existing fonts.

Therefore, it is wise to normalize the glyphs to use the PRC normalized form first. If the explicit form in the dictionary wish to be preserved, IVD should be registered after the normalized form is encoded.

00980 G_Z3951603: Consider normalize the right hand side to 兒 (U+5152) instead of 兜 (U+2B918).

G_Z3951603 龀
《古壮字字典》第395页左栏第6条
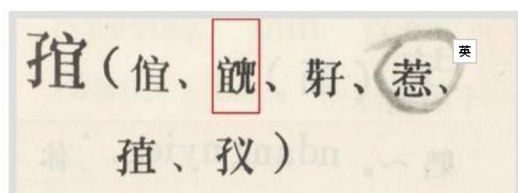
00224 G_Z2281101: Consider normalize to 觉 as 覚 is not a normalized form.

G_Z2281101 儥
《古壮字字典》第228页左栏第1条

gvo [kio¹]
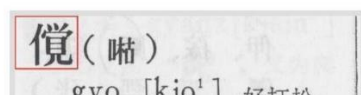
02271 G_Z2302301: same as above:

G_Z2302301 熄
《古壮字字典》第230页右栏第3条
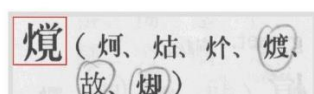
04111 G_Z1231301: The phonetic of the character is 忝 (tim1). It should be normalized to 忝.
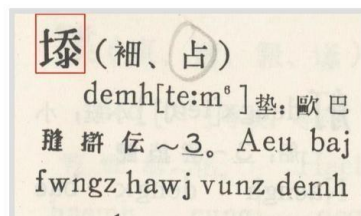
G_Z1231301 踮
《古壮字字典》第123页左栏第3条

踮（跕、点）

diem[ti:m¹] 提（脚）；
移动（脚步）：～町。
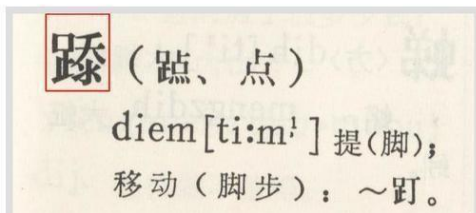
The confusion of 氺 and 心 for 忝 is common:

00739 G_Z1191201: same as above; duplicate of 塇 (U+2135E)

G_Z1191201 塇
《古壮字字典》第119页左栏第2条

塇（裧、占）

demh[te:m⁶] 垫：歐巴
雅 擤 伝 ～3。 Aeu baj
fwngz hawj vunz demh

03110 G_Z4442201: 斉 is not the PRC normative form. Consider normalize to 齐.

G_Z4442201 箭

《古壮字字典》第444页右栏第2条

箭（筒、鰤、灄）
saez [θai²] 鱼筌〔喇叭
形的〕: 列 箙 庩 紃～。
Lieb duk ma san saez.
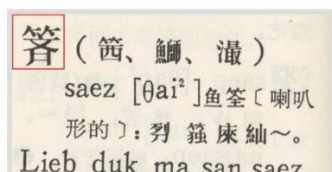
03187 G_Z0671501: 乖 is not PRC normative form. Consider normalize to 垂.

G_Z0671501 糎

《古壮字字典》第067页左栏第5条

糎（糙、糍、精、粃、
飿、粩、餹）
ceiz [ɕei²] 糍粑:
ⅲ～。guh ceiz. 做糍粑。
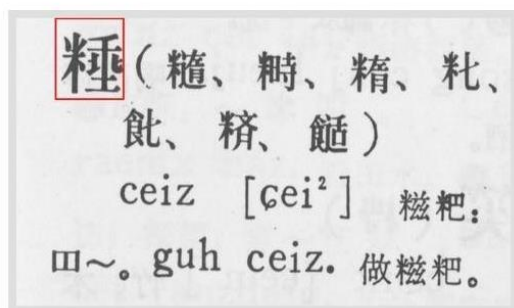
05030 G_Z3621104: 恶 is not the normative form. Consider normalize to 惡.
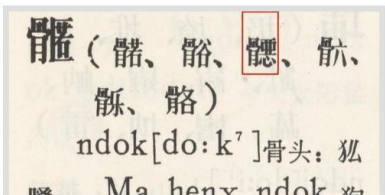
G_Z3621104 髝
《古壮字字典》第362页左栏第1条

03008 G_Z3271501: 宻 is not the PRC normative form. Consider normalize to 密.

G_Z3271501 稬
《古壮字字典》第327页左栏第5条
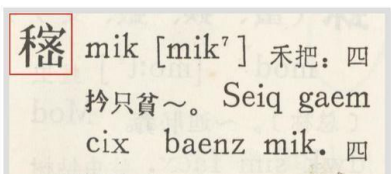
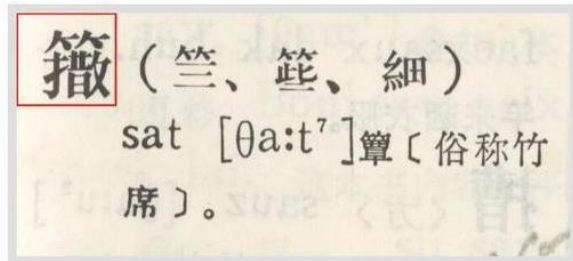

Please also refer to IRGN2154 ROK Normalization Rule 5-1:

5-1	宎(叉)	宓(必)	密-密 [A316;277c;6(丙)], 蜜-蜜 [b104;087a;1(丙)], 鄌 KC-05796 - 鄧 [A289;424d;9]	This type is various form of 必. 必 is from 丞→必 →尒 change to 叉 or 又.

03140 G_Z4491201: 籔 is not the PRC normative form. Consider normalize to 撒.

G_Z4491201 籔

《古壮字字典》第449页左栏第2条

籔（竺、篷、細）

sat [θa:t⁷]簟〔俗称竹席〕。

URL: https://github.com/hfhchan/irg-ws2015/issues/18

Issue 19

02315	爪 87.7		㒋 T13-2D5B				
	⬚⬚爪甫						
	點	N/A					

The left hand side is not Claw 爪 but 心 (忄). 怖 (U+6091) is already encoded.

According to Kangxi, 怖 is the variant of 怖, which is synonymous with 懼.

Action Item:

Unify/IVD to U+6091.

https://github.com/hfhchan/irg-ws2015/issues/19

Item 20

Owing to IRGN2211 Section B Item 3 "One-off corruptions found on tombstone carvings", the following characters should be rejected (or unified):

SN / Source / Treatment / Reason

02854 T13-2F48　　IVD 碑碑別字新編

02286 T13-2D55　　IVD 燦　　　碑別字新編

02270 TE-6F6B　　IVD 瞧　　　廣碑別字

02246 T13-2D4B　　IVD 照　　　偏類碑別字

02821 T13-2F42　　IVD 穎　　　碑別字新編

02812 T13-2F3F　　IVD 智　　　碑別字新編

02804 T13-2F3D　　IVD 矢　　　碑別字新編

02734 T13-2F29　　IVD 岂　　　廣碑別字

02750 T13-2F2F　　IVD 督　　　碑別字新編

02154 T13-2D32　　IVD 兊　　　廣碑別字

02088 T13-2D27　　IVD 澡　　　碑別字新編

02086 T13-2D24　　IVD 濱　　　碑別字新編

04289 T13-3138　　IVD 溯　　　廣碑別字

04299 T13-313C　　IVD 逮　　　廣碑別字

02073 T13-2D25　　IVD 淄　　　廣碑別字

02574 T13-2E49　　IVD 暴　　　偏類碑別字

02061 T13-2D21　　IVD 潰　　　偏類碑別字

01948 T13-2C54　　IVD 步　　　碑別字新編

02564 T13-2E48　　IVD 當　　　廣碑別字

02557 T13-2E42　　IVD 星　　　碑別字新編

01931 T13-2C4C　　IVD 氤碑別字新編

01933 T13-2C4D　　IVD 氳碑別字新編

01934 T13-2C4E　　IVD 氲碑別字新編

01929 T13-2C4A　　IVD 氣　　　偏類碑別字

02448 T13-2D79　　IVD 玉　　　偏類碑別字

02453 T13-2D7A   IVD 珍偏類碑別字

02437 T13-2D76   IVD 敵   碑別字新編

04003 T13-3128   IVD 敗   偏類碑別字

03968 T13-3124   IVD 短   碑別字新編

03803 T13-3075   IVD 福   碑別字新編

03777 T13-3074   IVD 礼   碑別字新編

03490 T13-304F   IVD 歸   偏類碑別字

03594 T13-3063   IVD 發   廣碑別字

03593 T13-3062   IVD 發   偏類碑別字

03367 T13-3044   IVD 孝   廣碑別字

03522 T13-3054   IVD 朝   偏類碑別字

01949 T13-2C53   IVD 沉廣碑別字

02642 T13-2E5F   IVD 癸   碑別字新編, missing stroke

02644 T13-2E60   IVD 癸   碑別字新編, protruding stroke

03610 T13-3068   IVD 發/羧   金石文字辨異

02036 T13-2C71   IVD 淑   金石文字辨異

04000 T13-3129   IVD 真   金石文字辨異

02845 T13-2F47   IVD 碑   金石文字辨異

02225 T13-2D45   IVD 婆   金石文字辨異

02097 T13-2D2B   IVD 演   金石文字辨異

02598 T13-2E51   IVD 病   金石文字辨異

Item 21

00198	人 9.9					佮 UTC-01573 (UK)	
	⿰亻⿱宀口口夊						
	橫	N/A					

Henry's Comment:

00198 UTC-01573 = 佮/佮, corrupted form of 侸U+20207 /佮U+20266

UK's Comment:

Disagree. UTC-01573 is a variant form of U+5BBF 宿, but U+20207 侸 is a variant form of U+5919 夙, so they are different characters and cannot be unified.

Henry's Further Comments:

夙 is often exchanged with 宿 in old Hanzi. Furthermore, the phonetic of 宿 is 佮 (stricter transcription -- 佮). The fact that the source says 00198 should be read as "U+5BBF 宿" does not necessarily contradict that 00198 can be unified with 侸U+20207 /佮U+20266.

Refer to the following source by Chinese University of Hong Kong (http://humanum.arts.cuhk.edu.hk/Lexis/lexi-mf/search.php?word=%E5%AE%BF):

Please note that the presence of the 宀 does not affect its meeting; the form with or without is found in Oracle Bone evidence and also Jianbo Wenzi.

Given that,

(1) 佋U+20207 and 㐼 U+20266 are variants of U+5BBF 宿 (per CUHK source)

(2) U+5BBF 宿 is variant with U+5919 夙 (per CUHK source)

(3) U+5919 夙 is variant of 佋U+20207 and 㐼U+20266 (provided by UK)

(4) 00198 is variant of U+5BBF 宿 (provided by UK)

(5) 00198 is virtually indistinguishable with U+20266 㐼, (and very likely referring to the same Oracle Bone/Old Hanzi glyph)

I conclude that 00198 should be unified with U+20266 㐼.

https://github.com/hfhchan/irg-ws2015/issues/21

Item 22

Possible Transcription Error:

00063 UTC-01316

IRGN2155CommentsToIRGN2107 (Chen Zhixiang)'s Comment:

4	UTC-01316	□□ 束	Fig. 4. A Concordance to Fascicle Three of the Inscriptions from the Yin Ruins (1985) p. 18 誺 488　壬戌卜(殸)貞：王塵陳？」	It is not a accurate transcription. should be transcribed to U+20094□ or U+2AA72 朿. So is U+2B85C 歸 束 is the transcription of

Henry's Comment

- WITHDRAW, reference IRGN2155CommentsToIRGN2107, OR
- unify to U+2B85C

UTC's Comment:

Disagree. Character is attested in two separate sources, and the right-hand side components are not unifiable.

Henry's Additional Comment:

The value of encoding erroneous transcriptions already identified by the Chinese experts should be justified. E.g. the "A Concordance to Fascicle Three of the Inscriptions from the Yin Ruins" is so academically significant that its error forms should be encoded as is, in similar respect to the Kangxi Dictionary and/or Hanyu Dazidian.

URL: https://github.com/hfhchan/irg-ws2015/issues/22

Item 23

There are a large number of characters where the U+2EBC has been replaced with the full 肉 radical. IRG should consider to allow encoding these characters as IVD as there are 500+ characters with U+2EBC. Variants using 肉 on the left are sufficiently rare in modern usage.

It is suggested that only IVD if 肉 on left; disunify if 肉 is at the bottom. Zhuang characters in particular should be studied on a case-by-case basis because 肉 may play a phonetic part instead of semantic part. Those with 肉 on the bottom have been included for completeness.

00282 USAT08988 IVD 舟: 肉 vs U+2EBC

02178 USAT08962 IVD 炙

02670 T13-2E6E IVD 胘

03390 USAT06462 IVD 肔: text indicates 乙 as phonetic; 肉 vs U+2EBC

03392 USAT09914 IVD 肋: 肉 vs U+2EBC

03393 USAT06266 IVD 肙: 肉 vs U+2EBC

03395 USAT08303 IVD 肥: 肉 vs U+2EBC

03397 USAT10232 IVD 肧: 肉 vs U+2EBC

03403 USAT05646 IVD 肴: 肉 vs U+2EBC, top 爻 (ref: 希~爷 // 04335)

03405 USAT90295 IVD 股/胑: 肉 vs U+2EBC

03407 USAT08919 IVD 育: 肉 vs U+2EBC

03411 USAT06030 IVD 背: 肉 vs U+2EBC

03414 USAT08746 IVD 胆: 肉 vs U+2EBC

03416 USAT90297 IVD 胜: 肉 vs U+2EBC

03419 USAT06375 IVD 肺: 肉 vs U+2EBC; IDS: ⿲肉市

03420 USAT10231 IVD 胎: 肉 vs U+2EBC

03422 USAT06361 IVD 胥: 肉 vs U+2EBC

03423 USAT08968 IVD 胞: 肉 vs U+2EBC

03426 USAT06768 IVD 胯: 肉 vs U+2EBC; (夲 ~ 夸 // 誇 ~ 誃 // 跨 ~ 踤 // 骻 ~ 躿 // 鮬 ~ 鮮 // 涝 ~ 涒);SC - 6

03437 USAT07202 IVD 腴: 肉 vs U+2EBC; SC = 9

03463 GHZR10093.03 IVD 臋

03470 USAT90298 IVD 臅

URL: https://github.com/hfhchan/irg-ws2015/issues/23