Universal Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N2803R2
Date: 2025-03-03

Source:	China
Author:	Tao Yang
Title:	Feedback on IRGN2788
Meeting:	IRG #64
Status:	Member Body Contribution
Actions required:	To be considered by IRG
Distribution:	IRG
Medium:	Electronic
Page:	6
Appendix:	None

Judith Chen has completed an inspection work on the document IRGN2788 and identified three kinds of data corrections that need to be revised. Here's a summary of the corrections:

1. Source Reference Value

China would agrees to Judith's suggestions except for item 1.6, for the other experts want to provide more effective methods to solve the G7 problem. Properties of U+4E85, U+5570 and U+7CA6 remain as G8-2F7C, G8-2F7D and G8-2F7B.

Here are the references which needed to be changed:

UCS	Glyph	Current Source	Revised Source
U+4FB4	俞	G3-327D	G5-313F
U+5DC2	嶌	G5-3F37	G3-3970
U+96DF	雧	G3-3970	G5-3F37
U+58AB	墫	G5-3722	GKX-0239.03
U+58FF	墫	GE-3541	G5-3722
U+4A9E	銮	G5-7768	G5-7767
U+23F7D	潛	GHZ-31737.08	GKX-0651.04
U+3ADA	舀	G3-4753	G3-4838
U+66F6	舀	G3-4838	G3-4753
U+4200	簡	G3-6429	GHZ-52973.13
U+5329	匩	GE-3270	GGT-00003
U+6FF2	瀫	GE-4037	GKX-0657.26
U+4E85	丄	G8-2F7B GU-04E85 GKX-0085.09	G8-2F7C
U+5570	啰	G8-2F7C GU-05570 GHZR-20691.09	G8-2F7D
U+7CA6	粦	G8-2F7A GU-07CA6 GKX-0909.01	G8-2F7B

2. Glyph Design

The G-glyph of U+5329 and U+6FF2 are correct, but they haven't been given the correct source. Please refer to the table above.

2.1 Explanation on the revision of U+5329 匩

The phonetic component 㞷 of U+5329 匩 is transcribed from Small Seal style glyph 㞷, which looks the same with the left component 㞷 of 封. However, this similarity is merely coincidental, arising during the stage of small seal script, the origins of them are quite different. 㞷 of U+5329 匩 originates from the Oracle style glyph  , which is composed of semantic component foot 止  (means the direction) and phonetic component king 王  (pronounced as wang). When the Oracle glyph transformed to Small Seal script, 止  and 王  were combined together, sharing a horizontal stroke in the middle  . So the best transcription of  㞷 should be designed like 㞷 with no breaks in the middle, and the best glyph of U+5329 should be 匩.

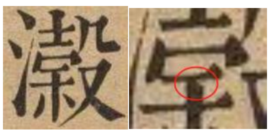
The glyph 匩 in GB/T 16500—1998 listed at 18-80 (0x3270) is different from what I expected, I would like to revise the glyph to 匩, and change the reference to GGT-00003.

There are already 5 encoded characters existing in UCS(㞷 U+37B7/𡉚U+2125A/𭖒U+2D592/𡴍U+21D0D/𰙝U+3065D) , I think it's necessary to encode 㞷 as a new component.

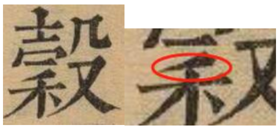
						
古俗字略	康熙字典	玉篇	字学正本	古韵溯原	字汇补	说文解字注

2.2 Explanation on the revision of U+6FF2

The glyph of U+6FF2 濲 doesn't need to be revised. It could be tell clearly that the middle part is ⿱青木 rather than ⿱青禾 in Kangxi Zidian.



While the first stroke of the character '禾' will be written as a clear ‘丿’.



So I would prefer not to change the glyph of U+6FF2 瀲 and change the reference to GKX-0657.26.

2.3 Hesitation about revision of U+9D56 鵖

It's obviously that the glyph writing in traditional style in GB 7589-87 follows the one in ancient literature.

As it has been reformed as  in UCS by following the normalization case of 既→既, the G-glyph of U+9D56 shows a quite different shape from the others 

, and the literature, published after the release of UCS, show the modified glyph form generally.

皀 is a component that does not require analogy, which retains its traditional form in the characters such as ⿷匚皀/⿱人皀/⿰氵皀/⿰犭皀/⿰木皀/⿰火皀/⿰氵⿸厂皀. So in this case, it should also remain unchanged 皀.

The only hesitation is that  has been already used commonly in Zoology for nowadays.

IDS	Glyph 1	Glyph 2	Glyph 3	Glyph 4	Glyph 5	Glyph 6
⿰皀鸟	 GB 7589-87	 鵖鴔(bīfú)		 说文解字点校本	 黄道周集	 清儒学案
⿰艮鸟	 鵖鴔，戴鵀	 礼记集解	 方以智全书	 *Oenanthe oenanthe* 穗鵖	 白尾黑鵖 白尾石鵖 冕鵖	 漠鵖 沙鵖 白顶鵖
⿰  鸟	 礼记注疏	 尔雅注疏				

2.4 need investigation to consider how to deal with the characters contain ⿰𠦝⿱𠂉/⿰𠦝⿱人

As Judith mentioned, Mr. Andrew West observed that 'there is no consistency in the use of ⿰𠦝⿱𠂉/⿰𠦝⿱人 in China characters', such is the fact.

There are 23 encoded characters which described by ⿰𠦝⿱人 in their IDS.

UCS code	glyph	UCS code	glyph	UCS code	glyph	UCS code	glyph
U+3CA6	㲦	U+7FF0	翰	U+24339	𤌹	U+29E7C	𩹼
U+4BA7	䮧	U+8792	螒	U+25010	𥀐	U+2A7FA	𪟺
U+501D	倝	U+203C9	𠏉	U+2524F	𥉏	U+2E2D6	𮋖
U+5E79	幹	U+224A8	𢒨	U+2648B	𦒋	U+2E37F	𮍿
U+65A1	斡	U+229E2	𢧢	U+27E73	𧹳	U+31160	𱅠
U+69A6	榦	U+23259	𣉙	U+29676	𩙶		

There are 9 encoded characters which described by ⿰𠦝⿱𠂉 in their IDS.

UCS code	glyph	UCS code	glyph	UCS code	glyph
U+4E7E	乾	U+20887	𠢇	U+293D1	𩏑
U+96D7	雗	U+2338D	𣎍	U+2C8BA	𬢺
U+9DBE	鶾	U+26A7B	𦩻	U+31DA1	𱶡

Meanwhile, there are more than 120 unencoded characters contain ⿰𠦝⿱𠂉/⿰𠦝⿱人, the glyph of them are indeed in a state of chaos as Andrew said.
It's not in a hurry to revise U+4BA7, U+9DBE and U+96D7, we would better verify all the glyph before making any further decisions.

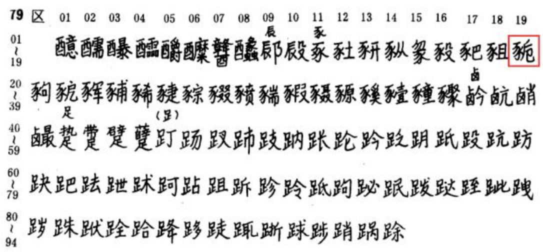
2.5 The following characters need their glyph revised:

UCS	GB 18030-2022 Glyph	Current UCS glyph	Revised UCS Glyph
U+5329	匩	匩	匩

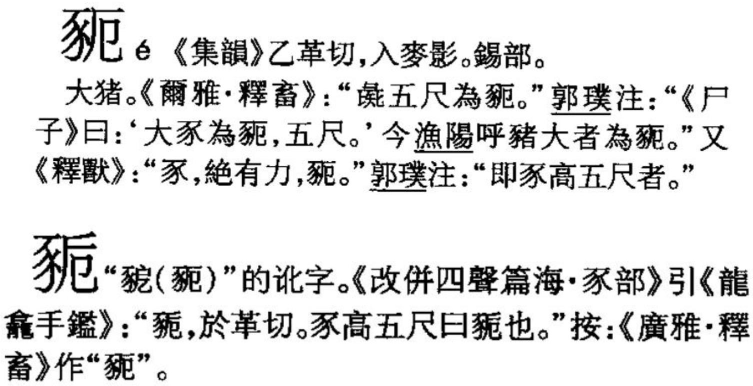
U+6F78	潸	潸	潸
U+23F7D	𣽽	𣽽	𣽽
U+3594	㖔	㖔	㖔
U+9D56	鵖	鵖	鵖
U+4748	䝈	䝈	䝈

3. Disunification

I think Judith's advise is dividing the glyph ⿰豕厄 to a new code point and change the G-glyph of U+4748 to ⿰豕戹.

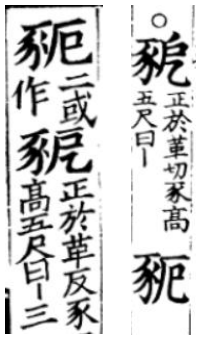
Current UCS	Current glyph		Recommended UCS
U+4748		 Figure 44: 79-19 (0x6F33) in GB 7589—87	U+xxxxx G3-6F33

Yes, we can make sure that the G-glyph ⿰豕厄 of U+4748 is a variant of 豟, so does the T/J/K-glyph ⿰豕戹. According to the evidence of Hanyu Dazidian, we can see that ⿰豕厄 comes from 龍龕手鑑. But the glyph in 2 versions of 龍龕手鑑 are ⿰豕戹 rather than ⿰豕厄, that means the note of ⿰豕厄 in Hanyu Dazidian is wrong.


豟 è 《集韻》乙革切，入麥影。錫部。
大猪。《爾雅·釋畜》："彘五尺為豟。"郭璞注："《尸子》曰：'大豕為豟，五尺。'今漁陽呼豬大者為豟。"又《釋獸》："豕，絶有力，豟。"郭璞注："即豕高五尺者。"

䝈 "豟(豟)"的讹字。《改併四聲篇海·豕部》引《龍龕手鑑》："䝈，於革切。豕高五尺曰䝈也。"按：《廣雅·釋畜》作"豟"。

Hanyu Dazidian 漢語大字典（第二版）



Longkan Shoujian 龍龕手鑑

While because I can find no more evidences for the glyph 豕厄 except GB 7589—87. The existence of this character is highly suspicious, it's not in hurry to encode 豕厄 before the new evidences appear.

I would prefer to modify the glyph of U+4748 only.

(EOF)