Doc Type:	Working Group Document
Title:	Request to add seven new UCVs[1]
Source:	Huáng Jùnliàng (黃俊亮)
Status:	Individual contribution
Action required:	To be considered by the IRG
Date:	March 6, 2024

1 留畱畱

1	2	3
留	畱	畱
U+7559	U+7571	U+24CA2
GHTJKPV	GTJKPV	T

Here are disunified examples. They share the readings and the meanings. Hence we suggest a level 1 UCV for 留畱畱.

U+7559 留		U+7571 畱		U+24A7E 畱	
U+3A28	㨨	U+3A45	㩅		
U+586F	塯	U+3667	㙧		
U+5EC7	廇	U+222BA	庮		
U+69B4	榴	U+6A4A	樊		
U+6E9C	溜	U+6F91	潏		
U+7460	瑠	U+74A2	瑢	U+24A7E	璢
U+7624	瘤			U+30906	瘤
U+905B	遛	U+285BB	遛		
U+9724	霤	U+29178	霤		
U+9DB9	鶹	U+2A173	鶹		
U+2254D	鹠	U+22532	徟		
U+22793	懰	U+2D7AC	懰 (⿰忄畱)		
U+24811	猶	U+2487C	猶 (⿰犭畱)		
U+256C5	褶	U+256FD	褶		
U+25837	稻	U+2588B	稲		
U+259E5	窗	U+25A0C	窗		
U+267A7	膤	U+2681D	膤		

[1] Sources of this document are available online: `https://github.com/JLHwung/IRGN2675`.

2 刭刉劫

	1	2	3
	刭	刉	劫
	U+5227	U+523C	U+52AB
	GHTJKPV	GHTKJP	GHTJKPV

Here are disunified examples. They share the readings and the meanings. Hence we suggest a level 1 UCV for 刭刉劫. Note that this UCV does not include U+5226 刦 because 刀 and 刂 are not unifiable.

U+5227 刭		U+523C 刉		U+52AB 劫	
U+272E4	𧋤			U+8710	蜐
		U+3958	㥘	U+393C	㤼
		U+20268	㔨	U+2CF86	㑰
		U+289F6	鉤	U+289B2	鉤
		U+2BF45	捌	U+22B31	拐

3 辛𢐌

	1	2
	辛	𢐌
	U+8F9B	U+2840C
	GHTJKPV	GTJ

Here are disunified examples. They share the readings and the meanings. Hence we suggest a level 1 UCV for 辛𢐌.

U+8F9B 辛		U+2840C 𢐌	
U+89EA	觧	U+278FF	觧
U+8F9F	辟	U+2E77B	辟
U+2B2C1	蟀	U+2E51D	蟀

4 昆昆

1	2
U+23B09	Unencoded
GTV	

Here are disunified examples. They share the readings and the meanings. In some cases 毚 can also be a variant of 毚, e.g. U+235F8 槀 is a variant of U+23535 槀. Therefore we suggest level 2 UCV for 昆昆. Note that the unification request for 鎚 with 鎚 was also brought up by Henry Chan during the WS2017 review[2].

U+23B09 昆		Unencoded 毚	
U+3521	勛	U+20892	勛
U+3BB0	槌	U+235FD	槌
U+5AB2	媲	U+5AD3	媲
U+818D	腈	U+2DA44	腈
U+8C94	貚	U+27D00	貚
U+939E	鎚	U+321DA	鎚
U+2A106	鷉	U+2A122	鷉

Although 毚 is not yet encoded, the proposed UCV does not imply that 毚, as a standalone character, should be unified to 昆. Quite the opposite I suggest we encode 毚 separately as it will be useful as an IDS component in the encoded characters mentioned above.

5 屑屑

1	2
U+5C51	U+3795
GHTJKP	GTP

---

[2]For more information, see `https://hc.jsecs.org/irg/ws2017/app/?id=04445`

Here are disunified examples. They share the readings and the meanings. Hence we suggest a level 1 UCV for 屑屑.

U+5C51 屑		U+3795 屑	
U+698D	榍	U+2354B	榍
U+2A64C	髉	U+2A647	髉
U+2A651	齰	U+2A646	齰

6   隸隸隸隸隸

	1	2	3	4	5
	U+96B7	U+96B8	U+28F7B	U+28F7E	U+2B559
	GTJKP	GHTJKP	GT	G	T

Here are disunified examples. They share the readings and the meanings. Hence we suggest a level 1 UCV for 隸隸隸隸隸. Note that this UCV can combine with UCV #435 to handle other variants, such as .

U+96B7 隸		U+96B8 隸		U+28F7B 隸		U+28F7E 隸		U+2B559 隸	
U+3611	嚇			U+2D2FA	嚇	U+2D31A	嚇		
U+22E00	擽			U+2D8BF	擽				
U+237CC	檪	U+2AD1B	檪						
U+240C0	瀮	U+2DCEC	瀮					U+240B0	瀮
		U+2533F	矑	U+25324	矙				
U+25DBE	籛	U+25DD7	籛						

7   寧寍

	1	2
	U+5BD5	U+21A34
	GHT	GHTP

There are currently no disunified examples, but they are all variants of 寧. When used as a component, they share the readings and the meanings. Hence we suggest a level 1 UCV for 寧寧. This UCV is similar to the UCV #267.

U+5BE7 寧		U+5BD5 寍		U+21A34 宁	
U+45FF	䗿	U+2E53E	䗿		
U+6FD8	濘	U+2C23F	濘		
U+8079	聹	U+265F0	聹		
U+85B4	薴	U+2C78A	薴		
U+944F	鑏	U+28B4B	鑏		
U+9E0B	鸋			U+2A162	鸋
U+261AD	繜	U+2615C	繜		
U+27C17	蹲	U+2B384	蹲		

Acknowledgement

(End of Document)
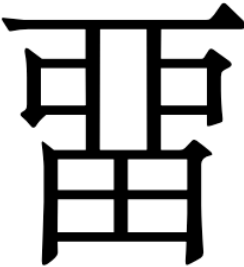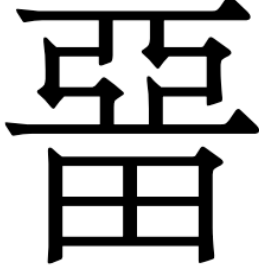
Feedback to IRGN2675 (Request to add seven new UCVs)

Author: Henry Chan
Date: 19 March 2024
Pages: 2

Feedback to discussion of Item 1

1	2	3
留	畱	𤲢
U+7559	U+7571	U+24CA2
GHTJKPV	GTJKPV	T

Per item 1, Wang Yifan raised an issue that shapes 1 and 2 should not be unified because they are represent different etymologies.

Per Shuowen, the characters which carry a modern shape of 卯 can either come from 卯 (夘) or 酉 (丣).

Per Shuowen, 留 actually comes from 丣 and is not related to 卯. Kangxi also uses the canonical form 畱 and calls 留 as the vulgar variant (俗字). However, in any modern orthography in East Asia, 留 is the predominant form if not the sole standard form.

As far as 留 is concerned, it doesn't really matter if it comes from 卯 (夘) or 酉 (丣). When used as a single character or as a phonetic component, 留 or 畱 is the exact same character. As it component there is no need to disunify them, and the character is already separately coded in the URO.

Therefore, I support Huang Junliang's suggestion to add it to the UCV so they can be unified when they are used as a component.

Feedback to Item 7

1	2	3	4	5
隸	隷	𨽻	𨽾	𫕙
U+96B7	U+96B8	U+28F7B	U+28F7E	U+2B559
GTJKP	GHTJKP	GT	G	T

I suggest putting "隸" as the first form, as all the others are usually recognised as the variant forms.

SAT Feedback to
"Request to add seven new UCVs" (IRGN2675)

<div align="right">Date: 2024-03-20</div>

Regarding the proposed UCV No. 1 in IRGN2675:

We believe that the glyph No. 1 and glyphs No. 2-3 (or, by extension, 留/畱 derived shapes) should not be made unifiable in principle.

The origin of the character 畱 can be traced back to 說文解字: "畱：止也。从田、丣聲。", where 說文解字 also state that 丣 is an older variant of 酉[1], different from 夘, which is original shape of 卯[2].

However, modern scholars generally believe that the ancestor of this character had always been attested with 卯 at the top in pre-Qin usage, and the glyph that has 丣 is an artifact introduced by 說文解字. Be noted though, that 留 is originally for a proper name at Oracle or Bronze script ages, and was rarely used in the current meaning until the Seal script stage, rather than being borrowed to write what are now 劉, 瘤, 流, etc.

Nevertheless, under the prestigious influence of 說文解字, 畱 (with its derivatives) has been received as the "canonical" shape of 留 for most of the history down to the Kangxi dictionary. Intellectuals do not only use 畱 as an "inkhorn" glyph interchangeably with 留, but sometimes

---

[1] 古文酉, 从丣。丣爲春門, 萬物已出。丣爲秋門, 萬物已入。

[2] 卯本字。與丣字上畫連者有別。丣音酉。

intentionally contrasted with 留 (see the picture).

古文酉從丣丣為春門萬物巳出丣
為秋門萬物巳入從一卯閉門象也

冒也二月萬物冒地而出象開門之形故二月為天
門凡卯之屬皆從卯臣鍇曰二月陰不能制陽陽冒

雷 也 止
從田丣聲臣鍇曰
田猶土也里由反

字作眠目冥也　稽雷　上經霓反考聲云楷　夜反
從目民聲也　滯也說文留止也從　留　夕
百秪聲秋音雞古文作乱或作叶下力求反　上
訓與前同考聲久也說文田也從田丣聲今
經文變體作留或作留又作
留展轉訛也丣音酉亦音桺　謝法

With above, an important and also complicating factor with regard to UCV here is that 說文解字 did actually alter the phonetic component. It is thought that 卯 had pronunciation *mruʔ and 酉 *luʔ[3] in the Old Chinese (i.e. around the time of 詩經), and they were eligible to 通假 (i.e. being options to be the phonetic indicator of the same word). Given that 說文解字 continues to use 丣 for 柳 or 聊, reanalysis of 留 as 🗌丣田 by 說文解字 strongly suggests that the phonetic system at 許愼 Xu Shen's period no more favors 卯 for 留's pronunciation as much as 酉. The fact effectively establishes 畱 a variant with alternative component (異構字) rather than a pure orthographic variant (異寫字) of 留, similar to what 綫 is to 線, or 仙 to 僊.

As far as we find, the current UCV list has few entries exhibiting a similar circumstance. Most are either difference in:

- component(s) which do not carry semantic or phonetic role; or
- component(s) whose cognacy with whichever option is unambiguous if the context given

However, of course, difference in a component which does not contribute to semantic distinction in isolation is not automatically marked as unification candidate, because we also recognize the principle of difference in actual component not to be unified. What we find could be counted as precedents of 留/畱 case might be UCV #183 (毎/每) and NUCV #294 (厤/麻), but the former is largely resulted from how late the 毋/母 distinction is fixated; and the latter is more related to simplified vs. traditional distinction. Moreover, in any case, 留/畱 retains quite dissimilar shapes in modern style as well as parallel usage until relatively recent times.

And as a tangential discussion, we think we should be more careful of pressing forward criteria solely based on holistic similarities, which could be in turn superficial. As far as I can remember, we have been already experienced several pitfalls in controversial cases such as [WS2021-02899 (🗌宀悤)](), [WS2021-01842 (🗌敄木)](), [WS2015-03555 (🗌艹仑)](), etc.

(End of document)

---

[3] Also *N-ruʔ according to Baxter and Sagart.