

Title: CJK & Unihan Working Group Recommendations for UTC #179 Meeting

Author: Ken Lunde

Date: 2024-04-12

This document provides the CJK & Unihan Working Group recommendations for the UTC #179 meeting, based on a meeting that took place on 2024-04-05 from 6 to 8PM PDT, which was attended by Yi Bai, Eiso Chan, Peter Constable, Koji Ishii, Kushim Jiang, John Knightley, Steven Loomis, Ken Lunde, William Nelson, CheonHyeong Sim, Stephan Hyeonjun Stiller, Harald Tveiten, Dennis Walters, Xieyang Wang, Yifán Wáng, and Andrew West via [Zoom](#). The Chair of the [CJK & Unihan Working Group](#), Ken Lunde, chaired the meeting.

The CJK & Unihan Working Group reviewed public feedback and documents that were received since the UTC #178 meeting took place. Comments are marked in **green**, and Recommendations to the UTC are marked in **red**. A “plain text” Copy&Paste-ready version of the Recommendations is also provided as the PDF attachment named *recommendations-20240405.txt*.

PRI #483 Feedback

The 11 items of [PRI #483](#) (*Proposed Update UAX #38*) feedback that are in this section were reviewed and discussed by the WG, and for convenience, the following data files, which are PDF attachments, include the recommended Unihan database additions and changes based on them, along with the recommended *CJKRadicals.txt* data file addition:

- *Unihan-additions-20240405.txt*
- *Unihan-changes-20240405.txt*
- *Unihan_IRGSources-changes-20240405.txt*
- *CJKRadicals-20240405.txt*
- *kMandarin-additions-20240405.txt*
- *kMandarin-changes-20240405.txt*

With the exception of the proposed *kMandarin* property value additions and changes, the recommended Unihan database additions and changes, along with the recommended *CJKRadicals.txt* data file addition, are also shown inline, following the **Recommendations**.

01) ID20240118004124

Date/Time: Thu Jan 18 00:41:24 CST 2024

ReportID: ID20240118004124

Name: Robin Leroy

Report Type: Public Review Issue

Opt Subject: **483**

In Unicode Versions 13.0 through 15.1, two characters have a kHanyuPinyin value but no kMandarin value:

U+228F5 𠄎 (kHanyuPinyin=42364.160:chú since Unicode 13.0)

U+2574C 禾 (kHanyuPinyin=42588.020:ji since Unicode 12.0)

These two characters should have the corresponding kMandarin values, thus

U+228F5 𠄎 should have kMandarin=chú, and

U+2574C 禾 should have kMandarin=ji.

Comments: The WG reviewed and discussed this feedback, and agreed that these would be appropriate *kMandarin* property value additions. As usual, these *kMandarin* property value additions will need to be checked by the CLDR-TC prior to the UTC accepting them.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Sharon Chen, CJK: Ask the CLDR-TC to check the proposed kMandarin property value additions, then report back to the UTC. See PRI #483 feedback [ID20240118004124] and Section 01 of document L2/24-067.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 01 of document L2/24-067. See PRI #483 feedback [ID20240118004124].

02) ID20240208072227

Date/Time: Thu Feb 08 07:22:27 CST 2024

ReportID: ID20240208072227

Name: Robin Leroy

Report Type: Public Review Issue

Opt Subject: **483**

While playing with invariant tests as part of an old action item (151-A143) now assigned to Josh Hadley and me, I noticed that the kTotalStrokes value of the Equivalent_Unified_Ideograph value of CJK strokes is not always 1; it is 2 for CJK STROKE HZZZG 𠄎, whose Equivalent_Unified_Ideograph is 𠄎.

Looking at the chart glyphs, and comparing with 𠄎 which is definitely meant to have a 𠄎, the G- and T-source glyphs look consistent with 𠄎, and at least the G-source glyph looks to me like it has one stroke. Should the kTotalStrokes value of 𠄎 be 1? Or perhaps 1 2? Or is the G-source glyph wrong? If <https://www.zdic.net/hans/%F0%A0%84%8E> and <https://www.zdic.net/hant/%F0%A0%84%8E> are to be believed, that would point towards the G-source glyph being right, and kTotalStrokes=1 2 (and kRSUnicode=5.0 6.1 I suppose).

While discussing this with a friend who is familiar with zh-Hant-TW, the kTotalStrokes value of 𠄎 also came up. While it has 2 strokes in zh-Hans-CN, this radical has three strokes in TW, see <https://dict.revised.moe.edu.tw/dictView.jsp?ID=11124>. Should its kTotalStrokes value be 2 3 (and likewise should the kRSUnicode=54.n characters have kTotalStrokes=n+2 n+3)?

I asked the above questions to Ken Lunde, who replied as follows:

- > What you discovered is the tip of the proverbial iceberg. See document
- > L2/22-201. This is a systemic issue, and our plan is to address it en masse,

- > not one-by-one. With that said, please submit feedback through the Contact
- > Form about this, and the CJK & Unihan Group will discuss. The IDS of U+2010E
- > 𠃉 does not lend itself to programmatic correction, so I would prefer to
- > handle this particular case separately.

Comments: The WG reviewed and discussed this feedback, and noted that the IRG counts U+4E43 乃 as having two strokes, so it would be correct to state that U+2010E 𠃉 should have only one stroke, but 2 should be kept as one of its *kTotalStrokes* property values. The WG also agreed that adding 5.0 to its *kRSUnicode* property value makes sense. About U+5EF4 𠃊, the IRG counts it as having three strokes, and three strokes is its official Kangxi Radical stroke count, so the WG recommends that its *kTotalStrokes* property value be changed to 3. There is a long-term effort underway to systematically correct *kTotalStrokes* property values, so the WG does not recommend at this time any adjustments to *kTotalStrokes* property values for ideographs that include 𠃊 as a component, but correcting the property value of the prime instance, U+5EF4 𠃊, is an important first step.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the proposal to change the informative *kRSUnicode* property value of U+2010E and the informative *kTotalStrokes* property values of U+5EF4 and U+2010E, based on PRI #483 feedback [ID20240208072227] and as amended in Section 02 of document L2/24-067, for Unicode Version 16.0.

Action Item for Michel Saignard, CJK: Change the *kRSUnicode* property value of U+2010E and the *kTotalStrokes* property values of U+5EF4 and U+2010E in ISO/IEC 10646's CJKSrc.txt data file, based on PRI #483 feedback [ID20240208072227] and as amended in Section 02 of document L2/24-067, and provide to Ken Lunde an updated Unihan16.txt data file, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Use the updated Unihan16.txt data file to update the Unihan database to change the *kRSUnicode* property value of U+2010E and the *kTotalStrokes* property values of U+5EF4 and U+2010E, based on PRI #483 feedback [ID20240208072227] and as amended in Section 02 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 02 of document L2/24-067. See PRI #483 feedback [ID20240208072227].

Current:

U+5EF4	<i>kTotalStrokes</i>	2
U+2010E	<i>kRSUnicode</i>	6.1
U+2010E	<i>kTotalStrokes</i>	2

Adjusted:

U+5EF4	<i>kTotalStrokes</i>	3
U+2010E	<i>kRSUnicode</i>	6.1 5.0
U+2010E	<i>kTotalStrokes</i>	1 2

03) ID20240215045359

Date/Time: Thu Feb 15 04:53:59 CST 2024

ReportID: ID20240215045359

Name: Ryusei Yamaguchi

Report Type: Public Review Issue

Opt Subject: **483**

The proposed change to the delimiter for the *kDefinition* property was initially suggested in Report ID: ID20231214115907, subsequently recorded in L2/24-012, and has been incorporated into the draft. However, using semicolons within the *kDefinition* property appears more akin to traditional dictionary punctuation rather than a delimiter for an array of strings. This raises several concerns:

1. Space After Semicolons: The current format for *kDefinition* values, such as "engraving tool; carve, engrave" for U+93B8 鑿, implies a JSON interpretation as ["engraving tool", " carve, engrave"], leaving an inappropriate trailing space.
2. General Descriptions: Some *kDefinition* entries provide general character properties unrelated to definitions per se, including language or variant information, e.g.: "(J) non-standard form of U+559C 喜, to like, love, enjoy; a joyful thing" for U+3402 喜. Descriptions for U+3421 𪗇 and U+342E 𪗞 include similar issues.
3. Semicolons Within Parentheses: Certain *kDefinition* values feature semicolons inside parentheses, which complicates any straightforward split by semicolons. This issue is evident in entries such as: For U+3A97 𪗇: "(a dialect) to open (a parcel; a bundle or a package); to unroll (a scroll, etc.); (Cant.) to rest, catch one's breath", where semicolons serve multiple roles within and outside parentheses. For U+25462 𪗇: "boron (obsolete; see U+787C 硼)", demonstrating how semicolons within parentheses can denote additional information rather than separate definitions.

Given these issues, it's recommended to reconsider or revert the proposed delimiter change.

For reference:

ID20231214115907: <https://www.unicode.org/review/pri483/feedback.html#ID20231214115907>

L2/24-012: <https://www.unicode.org/L2/L2024/24012-cjk-unihan-group-utc178.pdf>

Comments: The WG reviewed and discussed this feedback, and based on the issues that were pointed out, recommends that the *kDefinition* property syntax change be reverted and that the new “semicolon” delimiter be removed from PRI #483 (*Proposed Update UAX #38*). See PRI #483 feedback [ID20231214115907](https://www.unicode.org/review/pri483/feedback.html#ID20231214115907) that was discussed during the UTC #178 meeting. The WG also noted that this provisional property covers over 23,000 ideographs, so modifying them to resolve these issues would be an insurmountable task. Also, to be clear, reverting this syntax change does not require any *kDefinition* property values to be reverted or otherwise changed. This property should therefore continue to be free-form, which the WG believes was the original intention.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Ken Lunde, CJK: Update PRI #483 (Proposed Update UAX #38) to revert the Delimiter and Syntax change of the provisional kDefinition property, and to remove the description of the new “semicolon” Delimiter from Section 4.1, based on PRI #483 feedback [ID20240215045359] and Section 03 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 03 of document L2/24-067. See PRI #483 feedback [ID20240215045359].

Action Item for Ken Lunde, CJK: Convey to Mark Davis the CJK & Unihan Working Group comments in Section 03 of document L2/24-067. See PRI #483 feedback [ID20231214115907] and [ID20240215045359].

04) ID20240219000809

Date/Time: Mon Feb 19 00:08:09 CST 2024

ReportID: ID20240219000809

Name: Judith Chen

Report Type: Public Review Issue

Opt Subject: **483**

The kMandarin values of some characters may be erroneous.

- U+99CD 駢: péi -> pēng

As UAX #38 states, "(kMandarin is) The most customary pīnyīn reading for this ideograph. When there are two values, then the first is preferred for zh-Hans (CN) and the second is preferred for zh-Hant (TW). When there is only one value, it is appropriate for both." However, the preferred reading in Taiwan of U+99CD 駢 should be pēng, according to 重編國語辭典修訂本 (Revised Mandarin Chinese Dictionary) published by Taiwan's Ministry of Education and CNS11643 中文全字庫 website managed by Taiwan's Ministry of Digital Affairs. Besides, GB/Z 40637-2021 古籍印刷通用字规范字形表 (Standard glyph list of generally used Chinese characters for ancient books publishing) released by China's State Administration for Market Regulation also suggests that pēng is the preferred reading in PRC (p. 128), for the standard claims that "本表只收常用音" (the list records common readings only) (p. 230).

- U+42F3 緒: běi -> bèi

汉语大字典 (2nd ed) regards it as a variant of U+8919 緒 and gives bèi as the only reading (p. 3652). CNS11643 gives bèi, too.

- U+4876 鞞: bèi -> pì

GB/Z 40637-2021 (p. 130) and CNS11643 both give pì as the preferred reading.

- U+4A06 鰲: bí -> bì

汉语大字典 (2nd ed) regards it as a variant of U+9DE9 驚 (p. 4421), suggesting that its reading should be *bì*. CNS11643 gives *bì*, too.

- U+4D44 甹: *bí* -> *fēng*

汉语大字典 (2nd ed) regards it as a variant of U+9EB7 甹 (p. 4910), suggesting that its reading should be *fēng*. CNS11643 gives *fēng*, too.

Reference:

- GB/Z 40637-2021: <https://openstd.samr.gov.cn/bz/gk/gb/newGbInfo?hcno=52E2DE28D439C1937EE09AE4B5AA615B>

- 駢 in 重編國語辭典修訂本: <https://dict.revised.moe.edu.tw/dictView.jsp?ID=770&word=%E9%A7%8D>

- 駢 in CNS11643: <https://www.cns11643.gov.tw/wordView.jsp?ID=152944>

- 緝 in CNS11643: <https://www.cns11643.gov.tw/wordView.jsp?ID=216618>

- 鞞 in CNS11643: <https://www.cns11643.gov.tw/wordView.jsp?ID=287841>

- 瞿 in CNS11643: <https://www.cns11643.gov.tw/wordView.jsp?ID=219990>

- 甹 in CNS11643: <https://www.cns11643.gov.tw/wordView.jsp?ID=221497>

Comments: The WG reviewed and discussed this feedback, and agreed that these would be appropriate *kMandarin* property value changes. As usual, these *kMandarin* property value changes will need to be checked by the CLDR-TC prior to the UTC accepting them.

For reference, the following were the *kMandarin* property values for these ideographs in Unicode Version 6.0 (2010):

U+42F3	kMandarin	BEI4 MI4
U+4876	kMandarin	BEI4 PI4
U+4A06	kMandarin	BI2 BI4 XU4
U+4D44	kMandarin	BI2 B02 FENG4 PENG3
U+99CD	kMandarin	PENG1

They were changed to their current property values in Unicode Version 6.1 (2012):

U+42F3	kMandarin	běi
U+4876	kMandarin	bèi
U+4A06	kMandarin	bí
U+4D44	kMandarin	bí
U+99CD	kMandarin	péi

The proposed property value changes are as follows:

U+42F3	kMandarin	bèi
U+4876	kMandarin	pì
U+4A06	kMandarin	bì
U+4D44	kMandarin	fēng
U+99CD	kMandarin	pēng

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Sharon Chen, CJK: Ask the CLDR-TC to check the proposed *kMandarin* property value changes, then report back to the UTC. See PRI #483 feedback [ID20240219000809] and Section 04 of document L2/24-067.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 04 of document L2/24-067. See PRI #483 feedback [ID20240219000809].

05) ID20240222004423

Date/Time: Thu Feb 22 00:44:23 CST 2024

ReportID: ID20240222004423

Name: Eiso Chan

Report Type: Public Review Issue

Opt Subject: **483**

In kMandarin property values, there are 152 entries without the tones. In general, there must be the tones for the Putonghua reading of one single Hanzi (单字音), but some modern dictionaries (like 《现代汉语词典》) use the so-called fifth tone (aka 轻声) for some characters very discreetly, such as 们 reads as ·men not men. On the other hand, U+275C8 嶸 reads as n, it must be wrong.

Suggestions:

1. Re-check the Putonghua readings for these 152 Hanzi.
2. If the modern dictionaries use the so-called fifth tone for one Hanzi, we can accept the form as ·xxx (like ·men); if not, we should modify to the common forms.
3. Remove some entries if there are no acceptable Putonghua readings.

These 152 entries are shown as below.

U+35D1	bai	擘
U+37F7	da	嗒
U+4E48	me	么
U+4E5B	ya	ㄚ
U+4E7C	cui	攄
U+4E86	le	了
U+4E87	ma	𠂇
U+4EAA	ye	𠂇
U+4EEC	men	们
U+4F2C	ze	𠂇
U+4F66	shi	佑
U+4FA4	ta	俤
U+5011	men	們
U+516F	han	𠂇
U+5319	shi	匙
U+5417	ma	吗
U+5427	ba	吧
U+5440	ya	呀

U+5457	bei	呗
U+545A	wen	叟
U+5462	ne	呢
U+5497	zuo	叻
U+549C	ta	咤
U+54C7	wa	哇
U+54D8	xing	哢
U+5504	bei	唄
U+5525	lang	呤
U+554A	a	啊
U+5566	la	啦
U+55CE	ma	嗎
U+55E6	suo	嗦
U+561B	ma	嘛
U+561E	lei	嘞
U+5677	hm	噉
U+569C	me	嚙
U+56A1	hai	噍
U+56CE	zen	噍
U+56D6	lo	囉
U+578A	min	坨
U+57AF	da	圪
U+580F	fang	圪
U+58B6	da	撻
U+58ED	san	韦
U+5AF2	ma	嫫
U+5B50	zi	子
U+5C21	kun	焜
U+5DBF	ru	孺
U+5DEC	pu	奘
U+5DED	pu	奘
U+5F94	zhi	苈
U+5FC4	xin	卜
U+603D	mo	恸
U+6077	xiao	恸
U+6150	gong	憇
U+61F3	hui	懣
U+624C	shou	扌
U+62A3	yun	拘

U+63B5	ming	掄
U+63B9	meng	猛
U+63FC	beng	掇
U+6762	jiang	桤
U+67A0	zui	柞
U+6872	po	梲
U+6926	quan	榜
U+698B	chu	桴
U+6A75	san	撒
U+6A7A	jian	欄
U+6A7B	chu	楹
U+6AF5	jiao	樵
U+6B1F	guang	櫬
U+6C07	lu	魯
U+6C1E	bin	氤
U+6C35	shui	洎
U+6C62	tu	洎
U+6F48	zong	涿
U+6F9A	yu	灣
U+6FF9	me	溼
U+702E	ling	霖
U+709E	bian	炆
U+7140	wei	焜
U+7177	liang	焜
U+71DD	jing	焜
U+7220	ju	燿
U+7233	han	燿
U+74F2	wa	虺
U+7629	da	瘡
U+7666	me	瘻
U+7684	de	的
U+7740	zhe	着
U+7858	qing	硿
U+7A43	rong	榕
U+7A5D	zui	穉
U+7AA7	zhuo	罩
U+7B39	ti	篋
U+7BD2	shi	簞
U+7C17	liang	筵

U+7C2F	qi	筵
U+7C42	shi	篩
U+7C56	qian	籤
U+7C8C	yin	糲
U+7C8F	tai	糝
U+7D26	ba	紕
U+7DD5	qi	縉
U+7E4C	sha	纓
U+7E67	yun	纒
U+7E68	da	縫
U+7F3C	qi	飲
U+7F40	zhao	鈔
U+7FAA	yang	羴
U+810C	nin	胛
U+8126	de	臑
U+8279	cao	艹
U+8421	bo	苜
U+8457	zhe	著
U+848F	you	菑
U+84FF	xu	苻
U+8781	ban	蛄
U+87A9	tiao	蝮
U+87D0	chang	螳
U+88C4	xing	衿
U+88F3	shang	裳
U+8FF2	qu	馱
U+915C	fu	馱
U+933B	wu	鉞
U+9386	qian	鎗
U+93F1	zhang	鎗
U+93F2	qian	鎗
U+9466	xian	鑽
U+9596	shui	閑
U+97A1	la	鞞
U+990E	le	餽
U+9979	le	餽
U+9B98	dai	魷
U+9D64	jiao	鵯
U+9EB6	chi	麩

U+9EBC	me	麼
U+9EBF	mo	磨
U+20BBF	sa	屺
U+20C8E	fa	吡
U+20D68	de	得
U+21121	zhe	齧
U+2332D	hui	會
U+23424	jiu	相
U+23B36	ba	毘
U+23B37	ba	毘
U+24DDF	la	癩
U+25AFD	shi	筴
U+26570	duo	聃
U+266E8	lao	勝
U+275C8	n	嶸
U+292F7	la	鞞
U+29F87	hu	鯨

Comments: The WG reviewed and discussed this feedback, and agreed that the 152 *kMandarin* property values should be checked by the CLDR-TC.

Recommendations: The WG recommends to the UTC the following disposition ⇨

Action Item for Sharon Chen, CJK: Ask the CLDR-TC to check the 152 *kMandarin* property values, then report back to the UTC. See PRI #483 feedback [ID20240222004423] and Section 05 of document L2/24-067.

06) ID20240225211026

Date/Time: Sun Feb 25 21:10:26 CST 2024

ReportID: ID20240225211026

Name: Nathan G Glenn

Report Type: Error Report

Opt Subject: **Unihan**

The `kRSUnicode` field for the character 龜 has the apostrophe doubled, meaning that it does not match the syntax specification (this broke `unihan_etl`, which relied on the syntax specification for parsing).

Current value: 5.10 213''.0

Should be: 5.10 213'.0

Here's the page for it:

<https://www.unicode.org/cgi-bin/GetUnihanData.pl?codepoint=%E4%BA%80>

Comments: The chair of the WG discussed this feedback with the feedback submitter, which resulted in an acknowledgement that this was a well-documented change to the syntax of the informative *kRSUnicode* property. The chair also alerted the feedback submitter about another possible *kRSUnicode* syntax change that is covered in Section 10 of this document. Therefore, no UTC action is necessary.

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

07) ID20240311022911

Date/Time: Mon Mar 11 02:29:11 CDT 2024

ReportID: ID20240311022911

Name: Eiso Chan

Report Type: Public Review Issue

Opt Subject: **483**

The current *kCantonese* property value for U+282CD 𨮑 is wan1, but there are two values before 13.0.0: wan1 and wen1. In the daily life, we use wen1 not wan1. I also mentioned this issue when I proposed the corresponding simplified form in L2/24-056 with Liang Xin.

The current value is shown as below:

U+282CD *kCantonese* wan1

I suggest update as below:

U+282CD *kCantonese* wen1

Comments: The WG reviewed and discussed this feedback, and after some debate, agreed that the proposed *kCantonese* property value change is appropriate. The WG noted that the ideograph in question is used solely for representing the English-language loanword “van,” which explains why its Cantonese reading is unique. In addition, see Section 26 of this document for its simplified form, which is a proposed addition to UAX #45.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Ken Lunde, CJK: Update the UniHan database to change the *kCantonese* property value of U+282CD, based on PRI #483 feedback [ID20240311022911] and Section 07 of document L2/24-067, for Unicode Version 16.0.

Current:

U+282CD *kCantonese* wan1

Adjusted:

U+282CD *kCantonese* wen1

08) ID20240315202646

Date/Time: Fri Mar 15 20:26:46 CDT 2024

ReportID: ID20240315202646

Name: Paul Masson

Report Type: Error Report

Opt Subject: **Variants for U+9759 静**

This character is currently listed in the database as being its own simplified and traditional variant, both of which entries are patently incorrect. The correct traditional variant U+975C 靜 should of course remain in the database.

Comments: The WG reviewed and discussed this feedback, and noted that China does not consider these two ideographs to be in a simplified/traditional relationship. Therefore, no change is necessary.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 08 of document L2/24-067. See PRI #483 feedback [ID20240315202646].

09) ID20240325190110

Mon Mar 25 19:01:10 CDT 2024

ReportID: ID20240325190110

Name: Paul Masson

Report Type: Error Report

Opt Subject: **kTraditionalVariant for U+5729 圩**

This character is an alternate simplified form of U+589F 墟. This is not indicated in the database.

Comments: The chair of the WG discussed this feedback with the feedback submitter, and confirmed that the source of the information that led to this feedback is the same source as the provisional *kPhonetic* property. The WG agreed that the provisional *kSemanticVariant* property would be the appropriate way in which to capture the relationship between these two ideographs.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Ken Lunde, CJK: Update the Unihan database to add *kSemanticVariant* property values to U+5729 and U+589F, based on PRI #483 feedback [ID20240325190110] and as amended in Section 09 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 09 of document L2/24-067. See PRI #483 feedback [ID20240325190110].

New:

U+5729 kSemanticVariant U+589F<kPhonetic
U+589F kSemanticVariant U+5729<kPhonetic

10) ID20240328172102

Date/Time: Thu Mar 28 17:21:02 CDT 2024

ReportID: ID20240328172102

Name: Ken Lunde

Report Type: Public Review Issue

Opt Subject: **483**

One of the things that fell out of last week's IRG #62 meeting was the acknowledgement of a second non-simplified Chinese version of Radical 212 (aka dragon), which was encoded in its prime form in the Extension H block as U+31DE5 𪛗. The evidence in its IRG Working Set 2017 record clearly demonstrates the simplified/traditional relationship:

<https://hc.jsecs.org/irg/ws2017/app/index.php?id=03252>

This was discussed during the IRG meeting, and the IRG decided to append ".3" to identify a second non-Chinese simplified radical, meaning 212.3. The IRG appends ".0" to the radical number for traditional, ".1" for Chinese simplified, and ".2" for non-Chinese simplified. The latter two are equivalent to the use of the single and double apostrophe in the kRSUnicode property values. I propose that we add a triple apostrophe to handle a second non-Chinese simplified radical. Another side benefit is that the number that is appended to the Kangxi Radical number in IRG data will correspond to the number of apostrophes in the kRSUnicode property value.

In terms of other changes, the syntax of the kRSUnicode property would need to be updated, the text of UAX #38 would need to be updated to describe secondary non-Chinese simplified radicals, and the following line would need to be added to the UCD's CJKRadicals.txt data file (similar to the ones for 182'' and 208'' that lack a value for the second field):

212'''; ; 31DE5

The Extension H block also includes an ideograph that would benefit from a second kRSUnicode property value, so the following would be the two kRSUnicode property value changes:

U+31DE5 𪛗 kRSUnicode 117.6 → 212'''.0 117.6

U+31E22 𪛘 kRSUnicode 118.11 → 118.11 212'''.6

In addition, IRG Working Set 2021 includes four ideographs that would benefit from this second non-Chinese simplified radical, and their discussion records indicate that 212.3 will either be their primary or secondary radicals:

<https://hc.jsecs.org/irg/ws2021/app/index.php?id=00027> = 212.3 primary

<https://hc.jsecs.org/irg/ws2021/app/index.php?id=00880> = 212.3 secondary

<https://hc.jsecs.org/irg/ws2021/app/index.php?id=02082> = 212.3 secondary

<https://hc.jsecs.org/irg/ws2021/app/index.php?id=02595> = 212.3 secondary

That is all.

Comments: The WG reviewed and discussed this feedback, and agreed that accommodating a second non-Chinese simplified radical would be appropriate. With regard to the proposed *kRSUnicode* property value changes, the WG agreed that the first *kRSUnicode* property value of U+31DE5 should be **117.6**, not **211'''.0**. The chair of the WG also noted that the *RSIndex* app has already been updated to accommodate a second non-Chinese simplified radical.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the change to the syntax of the informative *kRSUnicode* property to accommodate a second non-Chinese simplified radical, the addition to the UCD's *CJKRadicals.txt* data file, and the changes to the informative *kRSUnicode* property values of U+31DE5 and U+31E22, based on PRI #483 feedback [ID20240328172102] and Section 10 of document L2/24-067, for Unicode Version 16.0.

Action Item for Markus Scherer, PAG: Update the UCD's *CJKRadicals.txt* data file, based on PRI #483 feedback [ID20240328172102] and Section 10 of document L2/24-067 and its attached *CJKRadicals-20240405.txt* data file, for Unicode Version 16.0.

Action Item for Michel Suignard, CJK: Change the *kRSUnicode* property values of U+31DE5 and U+31E22 in ISO/IEC 10646's *CJKSrc.txt* data file, based on PRI #483 feedback [ID20240328172102] and as amended in Section 10 of document L2/24-067, and provide to Ken Lunde an updated *Unihan16.txt* data file, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Use the updated *Unihan16.txt* data file to update the *Unihan* database to change the *kRSUnicode* property values of U+31DE5 and U+31E22, based on PRI #483 feedback [ID20240328172102] and as amended in Section 10 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Update PRI #483 (Proposed Update UAX #38) to revise the Syntax and Description of the informative *kRSUnicode* property, to update the sorting algorithm in Section 2.1.2, and to update Section 3.6 to reflect a second non-Chinese simplified radical, based on PRI #483 feedback [ID20240328172102] and Section 10 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Draft documentation for the Unicode Version 16.0 Beta and landing pages that specifies that parsers need to be aware of the presence of a second non-Chinese simplified radical in *kRSUnicode* property values and in the UCD's *CJKRadicals.txt* data file, based on PRI #483 feedback [ID20240328172102] and Section 10 of document L2/24-067, for Unicode Version 16.0.

New—*CJKRadicals.txt*:

212''' ; ; 31DE5

Current—*kRSUnicode* property syntax:

[1-9][0-9]{0,2}\{'0,2'\}\.-?[0-9]{1,2}

Adjusted—*kRSUnicode* property syntax:

[1-9][0-9]{0,2}\{'0,3'\}\.-?[0-9]{1,2}

Current:

U+31DE5	<i>kRSUnicode</i>	117.6
U+31E22	<i>kRSUnicode</i>	118.11

Adjusted:

U+31DE5	kRSUnicode	117.6	212''' .0
U+31E22	kRSUnicode	118.11	212''' .6

11) ID20240402063204

Date/Time: Tue Apr 02 06:32:04 CDT 2024

ReportID: ID20240402063204

Name: Judith Chen

Report Type: Public Review Issue

Opt Subject: **483**

In the current version of the UniHan database, U+8034 聒 has a kMandarin value of ‘yì’ and a kCantonese value of ‘ngat6’, while U+43B2 聒 has a kMandarin value of ‘zhé’, a kCantonese value of ‘zip3’, and a kDefinition value of ‘ear lobe; lobule’. However, these may not be correct.

As UAX #38 states, "(kMandarin is) The most customary pīnyīn reading for this ideograph. When there are two values, then the first is preferred for zh-Hans (CN) and the second is preferred for zh-Hant (TW). When there is only one value, it is appropriate for both." Therefore, we must consider their usage in both the PRC and Taiwan.

In the PRC, according to dictionaries like 《辞海》 and 《汉语大字典》，U+2652E 聒 means “耳垂” (‘ear lobe’) and is pronounced ‘zhé’ (陟葉切), while U+8034 聒 is used in the word “聒聒” which means “众多的声音” (‘many sounds’) and is pronounced ‘yì’ (魚乙切).

On the other hand, in Taiwan, according to 《異體字字典》 released by Taiwan’s Ministry of Education, the preferred reading of U+8034 聒 is ‘zhé’, while U+43B2 聒 has ‘yì’ as its preferred reading. Both characters mean “耳垂” (‘ear lobe’).

If we look at the kJapanese values of these characters, we may find them more confusing. U+8034 聒 has a kJapanese value of ‘チョウ ジョウ ニョウ’, the first coming from “陟葉切” while the latter two coming from “昵輒切”. U+43B2 聒 has a kJapanese value of ‘ギツ ゴチ’, which comes from “魚乙切”. U+2652E 聒 has a kJapanese value of ‘チョウ ジョウ’, similar to U+8034 聒.

There is no doubt that U+43B2 聒, composed of “耳” and “乙”, is a “形聲” (pictophonetic) character, and “乙” (the component on the right) reflects its pronunciation. Therefore, U+43B2 聒 should have a kMandarin value of ‘yì’ and a kCantonese value of ‘ngat6’. Since U+43B2 聒 in Taiwan can mean “耳垂” (‘ear lobe’), its kDefinition value can remain unchanged.

As for U+8034 聒, the thing is a little complicated. First, it is pronounced differently in the PRC and Taiwan, so its kMandarin value should include ‘yì’ and ‘zhé’. Moreover, its current H-source glyph is the same as the G-source glyph of U+2652E 聒, suggesting that its kCantonese value should also be ‘zip3’, the same as “輒”.

To summarise, I suggest that the UniHan database be changed as follows:

- U+8034 聒
 - kMandarin: yì -> yì zhé

- kCantonese: ngat6 -> zip3
- kDefinition: (empty) -> ear lobe; lobule
- U+43B2 聒
 - kMandarin: zhé -> yì
 - kCantonese: zip3 -> ngat6

That is all.

Comments: The WG reviewed and discussed this feedback, and agreed that Unihan database addition and changes are appropriate, but that the *kMandarin* property value changes will need to be checked by the CLDR-TC prior to the UTC accepting them.

Recommendations: The WG recommends to the UTC the following disposition ⇨

Action Item for Sharon Chen, CJK: Ask the CLDR-TC to check the proposed kMandarin property value changes, then report back to the UTC. See PRI #483 feedback [ID20240402063204] and Section 11 of document L2/24-067.

Action Item for Ken Lunde, CJK: Update the Unihan database to add a kDefinition property value to U+8034 and to change the kCantonese property values of U+43B2 and U+8034, based on PRI #483 feedback [ID20240402063204] and Section 11 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 11 of document L2/24-067. See PRI #483 feedback [ID20240402063204].

New:

U+8034	kDefinition	ear lobe; lobule
--------	-------------	------------------

Current:

U+43B2	kCantonese	zip3
U+8034	kCantonese	ngat6

Adjusted:

U+43B2	kCantonese	ngat6
U+8034	kCantonese	zip3

PRI #484 Feedback

PRI #484 (*Proposed Update UAX #50*) was re-issued shortly after the UTC #178 meeting, but no feedback was received prior to its closing date of 2024-04-02. Although the declared closing date has passed, its closing date should be extended to 2024-07-02 to accommodate feedback to be resolved during the UTC #180 meeting in July of 2024.

PRI #497 Feedback

The seven items of **PRI #497** (*Unicode 16.0 Alpha Review*) feedback that are in this section were reviewed and discussed by the WG, and for convenience, the following data files, which are

PDF attachments, include the recommended Unihan database changes and removals based on them:

- *Unihan-changes-20240405.txt*
- *Unihan-removals-20240405.txt*
- *Unihan_IRGSources-changes-20240405.txt*

The recommended Unihan database changes and removals are also shown inline, following the **Recommendations**.

12) ID20240217105436, ID20240327155804, ID20240328014549 & ID20240403090227

Date/Time: Sat Feb 17 10:54:36 CST 2024

ReportID: ID20240217105436

Name: Judith Chen

Report Type: Public Review Issue

Opt Subject: **497**

Considering that the KP-source glyph KP1-3413 of U+4E17 卐, which was added in Unicode 15.1, is identical to the G-source glyph GKX-0077.13 of U+2000D 卐 rather than other representative glyphs of U+4E17 卐, it might be a good idea if Unicode moved KP1-3413 from U+4E17 卐 to U+2000D 卐.

Date/Time: Wed Mar 27 15:58:04 CDT 2024

ReportID: ID20240327155804

Name: Andrew West

Report Type: Public Review Issue

Opt Subject: **497**

KP1-5653 is mapped to U+720B 燼 (𠄦火勳), but its glyph is actually the same as U+24455 燼 (𠄦火𠄦𠄦𠄦重灬力灬). Therefore KP1-5653 should be moved to U+24455.

Date/Time: Thu Mar 28 01:45:49 CDT 2024

ReportID: ID20240328014549

Name: Judith Chen

Report Type: Public Review Issue

Opt Subject: **497**

The KP-source glyph KP1-83F7 of U+96DF 蓊, which was added in Unicode 15.1, exactly matches all the representative glyphs of U+5DC2 蓊, rather than other representative glyphs of U+96DF 蓊. Therefore, I suggest moving KP1-83F7 from U+96DF 蓊 to U+5DC2 蓊.

Date/Time: Wed Apr 03 09:02:27 CDT 2024

ReportID: ID20240403090227

Name: Andrew West

Report Type: Public Review Issue

Opt Subject: **497**

Another incorrect KP mapping: KP1-4D4C (𠄦木𠄦糸育) currently maps to U+3BDE (𠄦木絹), but should map to U+23693 (𠄦木𠄦糸育).

Comments: The WG reviewed and discussed these four feedback items, and noted that all of the suggested *kIRG_KPSource* property value changes are recorded in document [L2/23-041](#) as so-called “suboptimal unifiable mappings.” Due to the lack of access to the actual DPRK standards, specifically KPS 10721 that corresponds to the “KP1” source prefix, the WG is reluctant to recommend any further changes to *kIRG_KPSource* property values. Instead, the author of [UTN #50](#) (*KP-Source Property Value History*) will add to the Unicode Technical Note a table that records the known suboptimal unifiable mappings for greater visibility.

Recommendations: The WG recommends to the UTC the following disposition ⇨

Action Item for Ken Lunde, CJK: Convey to the feedback submitters the CJK & Unihan Working Group comments in Section 12 of document L2/24-067. See PRI #497 feedback [ID20240217105436], [ID20240327155804], [ID20240328014549], and [ID20240403090227].

13) ID20240218072351

Date/Time: Sun Feb 18 07:23:51 CST 2024

ReportID: ID20240218072351

Name: Judith Chen

Report Type: Public Review Issue

Opt Subject: **497**

According to IRG N2276 by Jaemin Chung, there used to be some "pseudo-G8 characters" in UR0. The problem was partially solved in Unicode 13.0 by changing their *kIRG_GSource* values into self-referring GU sources. However, there are still two problems related to these "pseudo-G8 characters":

1. Characters like U+8980 𠄎, U+7CA6 𠄎, U+4E85 丨, U+5570 㗎 should not be changed to GU sources. Unlike other "pseudo-G8 characters", these characters do exist in GB 8565.2-88 and are, hence, expected to have a G8 source. The reason why N2276 also mentioned these characters is that their *kIRG_GSource* and *kGB8* values are wrong.

As Tianheng Shen wrote in IRG N2542, "It seems that these characters do not have a normal G-source, as if they are not used in the mainland of China." However, the fact is that U+5570 㗎 is listed as a level 1 character in China's 通用规范汉字表 (Table of General Standard Chinese Characters), which means that it is a common character in China.

Therefore, I suggest that the *kIRG_GSource* values of these four characters be changed as follows, based on N2276:

- U+8980 𠄎: GU-08980 -> G8-2F7A
- U+7CA6 𠄎: GU-07CA6 -> G8-2F7B
- U+4E85 丨: GU-04E85 -> G8-2F7C
- U+5570 㗎: GU-05570 -> G8-2F7D

2. N2276 also requested to modify the *kGB8* values of some characters. Unfortunately, this was not realised in Unicode 13.0.

I recommend removing the kGB8 values in the range 0883–0894, 1201–1294, and 1351–1394 from the Unihan database and correcting the kGB8 values of the following five characters, based on N2276:

- U+9B25 𠄎: 0893 -> 1589
- U+8980 𠄎: 1589 -> 1590
- U+7CA6 𠄎: 1590 -> 1591
- U+4E85 丨: 1591 -> 1592
- U+5570 𠄎: 1592 -> 1593

Comments: The WG reviewed and discussed this feedback, and agreed that the proposed property value changes are appropriate. For the normative *kIRG_GSource* property value changes, the WG contacted China to confirm that they agreed with the changes. The total number of *kGB8* property values to be removed is exactly 150.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the normative *kIRG_GSource* property value changes for U+4E85, U+5570, U+7CA6, and U+8980, based on PRI #497 feedback [ID20240218072351] and Section 13 of document L2/24-067, for Unicode Version 16.0.

Action Item for Michel Suignard, CJK: Change the *kIRG_GSource* property values of U+4E85, U+5570, U+7CA6, and U+8980 in ISO/IEC 10646’s CJKSrc.txt data file, based on PRI #497 feedback [ID20240218072351] and Section 13 of document L2/24-067, and provide to Ken Lunde an updated Unihan16.txt data file, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Use the updated Unihan16.txt data file to update the Unihan database to change the *kIRG_GSource* property values of U+4E85, U+5570, U+7CA6, and U+8980, based on PRI #497 feedback [ID20240218072351] and Section 13 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Update the Unihan database to change the *kGB8* property values of U+4E85, U+5570, U+7CA6, U+8980, and U+9B25, and to remove 150 *kGB8* property values, based on PRI #497 feedback [ID20240218072351] and Section 13 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 13 of document L2/24-067. See PRI #497 feedback [ID20240218072351].

Current:

U+4E85	kGB8	1591
U+4E85	kIRG_GSource	GU-04E85
U+5570	kGB8	1592
U+5570	kIRG_GSource	GU-05570
U+7CA6	kGB8	1590
U+7CA6	kIRG_GSource	GU-07CA6
U+8980	kGB8	1589
U+8980	kIRG_GSource	GU-08980
U+9B25	kGB8	0893

Adjusted:

U+4E85	kGB8	1592
U+4E85	kIRG_GSource	G8-2F7C
U+5570	kGB8	1593
U+5570	kIRG_GSource	G8-2F7D
U+7CA6	kGB8	1591
U+7CA6	kIRG_GSource	G8-2F7B
U+8980	kGB8	1590
U+8980	kIRG_GSource	G8-2F7A
U+9B25	kGB8	1589

Remove:

U+4E26	kGB8	1286
U+4EB8	kGB8	1202
U+4F3E	kGB8	1203
U+501E	kGB8	1204
U+5091	kGB8	1289
U+50C7	kGB8	1205
U+50CD	kGB8	1290
U+52B9	kGB8	1291
U+52F3	kGB8	1351
U+5338	kGB8	0890
U+53BE	kGB8	1201
U+540B	kGB8	0883
U+544E	kGB8	0884
U+5521	kGB8	0885
U+5557	kGB8	1293
U+5574	kGB8	1224
U+5586	kGB8	1362
U+55A4	kGB8	1225
U+5640	kGB8	1226
U+5684	kGB8	1227
U+577F	kGB8	1361
U+5848	kGB8	1217
U+590A	kGB8	0891
U+59B3	kGB8	1370
U+59DE	kGB8	1231
U+59FD	kGB8	1232
U+5A5E	kGB8	1233
U+5AAE	kGB8	1371
U+5C05	kGB8	1365
U+5D11	kGB8	1367
U+5D19	kGB8	1368
U+5D1F	kGB8	1228
U+5DD6	kGB8	1369
U+5E0B	kGB8	1366
U+5EA4	kGB8	1215
U+5EC8	kGB8	1359
U+5F07	kGB8	1220
U+60AE	kGB8	1358
U+6165	kGB8	1214
U+6245	kGB8	1239
U+624A	kGB8	1240
U+6294	kGB8	1221
U+647D	kGB8	1222
U+658F	kGB8	1373
U+65C2	kGB8	1394

U+6607	kGB8	1380
U+6673	kGB8	1382
U+668E	kGB8	1381
U+6767	kGB8	1247
U+68BF	kGB8	1379
U+691D	kGB8	1377
U+6A11	kGB8	1378
U+6AB5	kGB8	1248
U+6B38	kGB8	1258
U+6BA3	kGB8	1250
U+6C25	kGB8	1383
U+6C96	kGB8	1352
U+6C98	kGB8	1207
U+6CC3	kGB8	1209
U+6CDC	kGB8	1208
U+6D44	kGB8	1353
U+6D6C	kGB8	0886
U+6E1B	kGB8	1354
U+6E5D	kGB8	1210
U+6E67	kGB8	1355
U+6E96	kGB8	0889
U+6EA6	kGB8	1211
U+6EEB	kGB8	1212
U+6F82	kGB8	1356
U+6FA5	kGB8	1213
U+6FEC	kGB8	1357
U+709E	kGB8	1374
U+71CA	kGB8	1249
U+71CF	kGB8	1238
U+71FF	kGB8	1376
U+725A	kGB8	1223
U+72A8	kGB8	1255
U+72C9	kGB8	1229
U+739A	kGB8	1242
U+742B	kGB8	1243
U+7488	kGB8	1244
U+74AA	kGB8	1245
U+74D8	kGB8	1246
U+74E9	kGB8	0887
U+7526	kGB8	1285
U+7551	kGB8	1375
U+7652	kGB8	1384
U+76EA	kGB8	1386
U+770A	kGB8	1263
U+778B	kGB8	1385
U+797C	kGB8	1241
U+7B5C	kGB8	1273
U+7B87	kGB8	1390
U+7C50	kGB8	1391
U+7CF5	kGB8	1219
U+7E44	kGB8	1276
U+7EBB	kGB8	1234
U+7EE4	kGB8	1235
U+7EF9	kGB8	1236
U+7FD9	kGB8	1274

U+7FEF	kGB8	1275
U+8129	kGB8	1288
U+8152	kGB8	1257
U+8453	kGB8	1218
U+8460	kGB8	1364
U+849E	kGB8	1363
U+870E	kGB8	1272
U+8780	kGB8	1388
U+87A3	kGB8	1256
U+882D	kGB8	1389
U+8897	kGB8	1260
U+88AF	kGB8	1261
U+8955	kGB8	1262
U+897E	kGB8	0892
U+89D4	kGB8	1292
U+8BAC	kGB8	1287
U+8D5F	kGB8	1253
U+8E36	kGB8	1278
U+8E45	kGB8	1277
U+8E5C	kGB8	1280
U+8E62	kGB8	1279
U+8EAD	kGB8	1392
U+8F80	kGB8	1251
U+8F92	kGB8	1252
U+8FFA	kGB8	1360
U+9118	kGB8	1206
U+94C7	kGB8	1294
U+94DA	kGB8	1264
U+955A	kGB8	1265
U+9560	kGB8	1266
U+9575	kGB8	1393
U+9618	kGB8	1216
U+96BB	kGB8	0888
U+9778	kGB8	1281
U+9895	kGB8	1387
U+98D0	kGB8	1259
U+998C	kGB8	1230
U+9A94	kGB8	1372
U+9A99	kGB8	1237
U+9B25	kGB8	0893
U+9B31	kGB8	0894
U+9B36	kGB8	1254
U+9B46	kGB8	1282
U+9E24	kGB8	1267
U+9E40	kGB8	1268
U+9E50	kGB8	1269
U+9E5D	kGB8	1270
U+9E60	kGB8	1271
U+9F2B	kGB8	1283
U+9F41	kGB8	1284

14) ID20240303221223

Date/Time: Sun Mar 03 22:12:23 CST 2024

ReportID: ID20240303221223

Contact: eisoch@126.com

Name: Eiso Chan

Report Type: Public Review Issue

Opt Subject: 497

The J-Source glyphs for U+2011E 𠄎 (JMJ-030462), U+2011F 𠄏 (JMJ-030463) and U+20120 𠄐 (JMJ-030464) are designed for Sung/Ming style. The Moji Joho Kiban database pages show they are really Han characters/Kanji, so the J glyphs should be updated.

Comments: The WG reviewed and discussed this feedback, and recommends that the chair of the WG be assigned an action item to forward this feedback to the Japan NB for consideration. The WG also noted that according to 大漢和辞典, the usage of U+20120 𠄐 appears to be similar to that of U+16FE3 𠄑 OLD CHINESE ITERATION MARK. The excerpt below is from Volume 1, Page 453 of 大漢和辞典 that shows the entries for U+2011E (248), U+2011F (249), and U+20120 (250):

<p>【248】 シ、ヤウ 上(1-13)の古文。(正字通)通(1)古文上。</p> <p>【249】 カ 下(1-14)の古文。(正字通)古文下。</p> <p>【250】 同じ文字を重ねる記號。(容齋隨筆、元二之災)古書、字當再讀一者、即于上字之下、爲小二字、此字、當兩度言之、云云、今岐州石鼓銘、凡重言者、皆爲二字、明驗也。(通俗編、識餘)古鐘鼎文、子孫二字、皆不複書、周宣石鼓文、君子員獵員遊、雖四字相間、猶作員獵二、漢石經、改篆爲八</p>	<p>ける賦、曾鞏に於ける韻語、黃庭堅に於ける散語をいふ。(後山叢書)蘇明允不能詩、歐陽永叔不能賦、曾子固短韻語、黃魯直短散語。</p> <p>【一没食子酸】¹²²⁰ 鞣酸(9-42909-1)に同じ。</p> <p>【一餘堂叢書】¹²²¹ 叢書の名。清、師範編。其の目は、四六線起一卷(孫梅)、魏伯子雜說一卷(魏際瑞)、魏叔子外篇一卷(魏禧)、微波樹遺書一卷(孔繼涵)、習園藏稿一卷(余慶春)、鄂亭詩話一卷(磊河山房本)、切時政要一卷(韓學渾)、蜀游記一卷(沈)、楚詞達一卷(魯華)、瀛海虞衡志十三卷(檀萃)、課餘隨錄三卷(師範)。(叢書舉要、近代十)</p> <p>【二硫化二砷】¹²²² 砒素の硫化物。褐赤色の粉末。天然には、雞冠石として産す。顔料・煙火などに用ひる。雌黄。二硫化砷素。</p> <p>【二氯化碳醣】¹²²³ 無色の氣體。軍事上、毒瓦斯として用ひられる。フオスゲン。</p> <p>【二氯化乙硫醚】¹²²⁴ 芥子臭のある無色の液體。猛毒性を有し、糜爛性毒瓦斯として用ひられる。ザクローエチルサルファイド。イペリット。</p>	<p>分、如易之乾三、書之安二、亦如之、後漢書節傳、時遭二之災、章懷注、元二即元元也、蓋古人楷書、又有然者、二翁雜說、複語書二字、重二文也、升庵外集、二乃古文上字、言字同于上、省複書也、按、二說未定孰是、今人或書二、或書匕、各于舊說有合。(陔餘叢考、重字二點)凡重字、下者可作二畫、始於石鼓文重字、皆二畫也、後入襲之、因作二點。</p>
<p>【252】 𠄎 (集韻)雲俱切 𠄏 (正韻)休居切 𠄐 (詩集傳)音吁</p> <p>小 𠄎 ああ。ここに。發語の辭。もと𠄎(1-26)に作る。(説文)𠄎、於也、象氣之舒𠄎、从𠄎从一、一者其氣</p>	<p>【251】 𠄎 (集韻)丑玉切 𠄏 (集韻)株遇切 𠄐 (集韻)株遇切</p> <p>小 𠄎 止まる。たたずむ。(説文)𠄎、步止也、从二反イ、讀若畜。(左思、魏都賦)澤馬于阜。𠄎少しくあゆむ。(字彙)𠄎、小步也。𠄎右足のあゆみ。(正字通)𠄎、左步爲𠄎、右步爲𠄎、合之、則爲𠄎行字。(顏延之、楮白馬賦)纖驪接趾、秀騏齊𠄎。𠄎(姓氏尋源)イ、義通於躑躅、或爲𠄎行氏所改。</p> <p>解字 指事。イの反對の字形。イは股、脛、足の三つを指し、ふらふら歩くの意。故に𠄎は、ふらふら歩いて戻りとどまること。(説文通訓定聲)𠄎、步止也、从二反イ、指事、躑躅、即イ𠄎也。</p>	<p>一畫</p>

Also see the Moji Jōhō Kiban (文字情報基盤) database entries for U+2011E (MJ030462), U+2011F (MJ030463), and U+20120 (MJ030464). For reference, the code chart excerpts for U+2011E, U+2011F, and U+20120 as they appear in Unicode Version 16.0 Alpha are shown below:

2011E

二 7.0

————▲	————▲
————▲	————▲
GKX-0086.03	T6-2127

JMJ-030462

2011F

二 7.0

————▲	————▲
————▲	————▲
GKX-0086.04	T6-2128

JMJ-030463

20120

二 7.0

————▲	————
————▲	————
T6-2129	JMJ-030464

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Ken Lunde, CJK: Forward to the Japan NB the CJK & Unihan Working Group comments in Section 14 of document L2/24-067. See PRI #497 feedback [ID20240303221223].

15) ID20240310045725

Date/Time: Sun Mar 10 04:57:25 CDT 2024

ReportID: ID20240310045725

Name: Michel Mariani

Report Type: Public Review Issue

Opt Subject: **497**

In Version 16.0 ALPHA REVIEW of the Code Charts:

<https://www.unicode.org/Public/draft/UCD/charts/CodeCharts.pdf>

The V-Source glyph (V1-6C40) for U+99D5 駕 appears to be defective (incomplete horse component 馬); it is probably based on the glyph defined in the Vietnamese font Nom Na Tong v4.6, which has been corrected since v4.8 (currently v5.09).

Comments: The WG reviewed and discussed this feedback, and noted that this particular V-source representative glyph defect has been present in the code charts beginning from Unicode Version 14.0 (2021):

99D5	駕	駕	駕	駕	駕	駕	駕
馬 187.5	G1-3C5D	HB1-BE72	T1-6F7B	J0-326F	K0-4A3D	KP0-CDBC	V1-6C40

For comparison, the glyph for U+99D5 from *Nom Na Tong* Version 5.09 is shown below:

駕

The WG also noted that this relatively minor representative glyph defect does not rise to the level of requiring a glyph erratum to be issued.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Ken Lunde, CJK: Work with Vietnam to provide to Michel Suignard and the Charts Working Group an updated V-source font, based on PRI #497 feedback [ID20240310045725] and Section 15 of document L2/24-067 for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 15 of document L2/24-067. See PRI #497 feedback [ID20240310045725].

Public Feedback

The four items of public feedback from document [L2/24-063](#) that are in this section were reviewed and discussed by the WG, and for convenience, the following data files, which are PDF attachments, include the recommended Unihan database additions and changes based on them, along with the recommended UAX #45 data file changes:

- *Unihan-changes-20240405.txt*
- *Unihan_IRGSources-additions-20240405.txt*
- *USource-changes-20240405.txt*
- *kMandarin-additions-20240405.txt*

With the exception of the proposed *kMandarin* property value additions, the recommended Unihan database additions and changes, along with the recommended UAX #45 data file changes, are also shown inline, following the **Recommendations**.

16) ID20240112185925

Date/Time: Fri Jan 12 18:59:25 CST 2024

ReportID: ID20240112185925

Name: Paul Masson

Report Type: Error Report

Opt Subject: **kMandarin for U+26760 羸 and U+21FEA 𪛗**

Both of these characters have two different pronunciations, neither of which is listed for each one. It is unclear to me what the definitive pronunciations should be for each, but there should at least be some entry in the database.

Comments: The WG reviewed and discussed this feedback, and came up with the following *kMandarin* property values that the CLDR-TC will need to check prior to the UTC accepting them:

U+21FEA	kMandarin	liè
U+26760	kMandarin	luó

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Sharon Chen, CJK: Ask the CLDR-TC to check the proposed *kMandarin* property value additions, then report back to the UTC. See feedback [ID20240112185925] in document L2/24-063 and as amended in Section 16 of document L2/24-067.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & Unihan Working Group comments in Section 16 of document L2/24-067. See feedback [ID20240112185925] in document L2/24-063.

17) ID20240201123549

Date/Time: Thu Feb 01 12:35:49 CST 2024

ReportID: ID20240201123549

Name: Lee Collins

Report Type: Error Report

Opt Subject: **Unihan_Readings.txt**

Defintion of U+4AB3 is grammatically incorrect.

Current (v 15.0):

U+4AB3 kDefinition slanted face causing by the paralyzed of the facial nerve

Better:

U+4AB3 kDefinition slanted face caused by paralysis of the facial nerve

Comments: The WG reviewed and discussed this feedback, and while there was agreement that the *kDefinition* property value of U+4AB3 is grammatically incorrect, a slightly different wording is recommended.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Ken Lunde, CJK: Update the UniHan database to change the kDefinition property value of U+4AB3, based on feedback [ID20240201123549] in document L2/24-063 and as amended in Section 17 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the feedback submitter the CJK & UniHan Working Group comments in Section 17 of document L2/24-067. See feedback [ID20240201123549] in document L2/24-063.

Current:

U+4AB3 kDefinition slanted face caus**ing by the paralyzed of the fa-**
cial nerve

Adjusted:

U+4AB3 kDefinition slanted face caus**ed by paralysis of facial nerves**

18) ID20240308071114

Date/Time: Fri Mar 08 07:11:14 CST 2024

ReportID: ID20240308071114

Name: Ken Lunde

Report Type: Error Report

Opt Subject: **UAX #45**

While compiling the candidates for the UTC's submission for IRG Working Set 2024, which included preparing the metadata, I came across a small number of attributes that should be changed to the following UAX #45 records:

UTC-00358: First Stroke = 1; Total Strokes = 10

UTC-00373: First Stroke = 3; Total Strokes = 17

UTC-00792: Variant = U+8F66; First Stroke = 0; Total Strokes = 4; UTCDoc (add) = L2/21-149 1

UTC-00910: First Stroke = 5

UTC-01257: IDS = 𠄎 𠄏 𠄐

UTC-03154: IDS = 𠄑 𠄒

UTC-03247: IDS = 𠄓 𠄔

UTC-03248: UTCDoc = L2/21-149 2

UTC-03254: IDS = 𠄕 𠄖 𠄗 𠄘 𠄙 𠄚

UTC-03257: IDS = 𠄛 𠄜

UTC-03268: Stroke Count = 7; Total Strokes = 15; IDS = 𠄝 𠄞 𠄟 𠄠 𠄡 𠄢 𠄣

UTC-03273: IDS = 𠄤 𠄥

UTC-03277: IDS = 𠄦 𠄧

UTC-03279: IDS = 𠄨 𠄩

UTC-03374: IDS = 𠄪 𠄫 𠄬 𠄭 𠄮

UTC-03380: IDS = 𠄯 𠄰

UTC-03392: IDS = 𠄱 𠄲

UTC-03247;FutureWS;U+84AD;140.6;;𣎵𣎶;UTCDoc L2/21-101 9;JTI code を;10;5
 UTC-03248;FutureWS;U+807F;129.0;;𣎵一;UTCDoc L2/21-149 2;;6;0
 UTC-03254;FutureWS;;147.21;;𣎵𣎵見見𣎵𣎵見見;UTCDoc L2/22-009 3;;28;2
 UTC-03257;FutureWS;;85.7;;𣎵𣎵𣎵;UTCDoc L2/22-040 3;;11;2
 UTC-03268;FutureWS;;167.7;;𣎵𣎵金𣎵𣎵𣎵;UTCDoc L2/22-040 14;;15;1
 UTC-03273;FutureWS;;85.21;;𣎵𣎵雨𣎵;UTCDoc L2/22-040 19;;24;1
 UTC-03277;FutureWS;;75.7;;𣎵𣎵𣎵秋;UTCDoc L2/22-040 23;;11;1
 UTC-03279;FutureWS;;75.10;;𣎵𣎵木𣎵𣎵𣎵;UTCDoc L2/22-040 25;;14;4
 UTC-03374;FutureWS;;75.8;;𣎵𣎵木𣎵𣎵止𣎵𣎵𣎵八;UTCDoc L2/23-130 26;kJapanese しぶ;12;1
 UTC-03380;FutureWS;;140.14;;𣎵𣎵𣎵𣎵廣;UTCDoc L2/23-130 32;;17;4
 UTC-03392;FutureWS;;140.6;;𣎵𣎵𣎵𣎵𣎵;UTCDoc L2/23-139R 8;Japanese simplified form of U+85E4;10;3

19) ID20240316075059

Date/Time: Sat Mar 16 07:50:59 CDT 2024

ReportID: ID20240316075059

Name: Ken Lunde

Report Type: Error Report

Opt Subject: **UAX #45 USourceData.txt**

Per the following post, ✕ user @Kesuuko_0826 indicated that UAX #45 ideograph UTC-03332 is encoded in the Extension I block as U+2ECAC 𣎵:

https://twitter.com/Kesuuko_0826/status/1768903062461030446

Its record in the USourceData.txt data file should be changed to the following, and U+2ECAC 𣎵 can be horizontally extended to add UTC-03332 as a new U-source source reference:

UTC-03332;ExtI;U+2ECAC;75.4;;𣎵𣎵木太;UTCDoc L2/22-206 35;;8;1

That is all.

Comments: The WG reviewed and discussed this feedback, and agreed that U+2ECAC should be horizontally extended to add UTC-03332 as a new U-source source reference and that its UAX #45 record be updated accordingly.

The table below constitutes a proposal to horizontally-extend U+2ECAC by adding UTC-03332 as a new U-source source reference:

Code Point	Code Chart Excerpt—Version 15.1	U-Source	
		Source Reference	Representative Glyph
U+2ECAC	2ECAC 𣎵 木 75.4 GIDC23-189	UTC-03332	𣎵

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the proposal to add UTC-03332 as a new normative kIRG_USource property value for U+2ECAC, based on feedback [ID20240316075059] in document L2/24-063 and Section 19 of document L2/24-067, for Unicode Version 16.0.

Action Item for Michel Suignard, CJK: Add UTC-03332 as a new `kIRG_USource` property value for U+2ECAC in ISO/IEC 10646's CJKSrc.txt data file, based on feedback [ID20240316075059] in document L2/24-063 and Section 19 of document L2/24-067, and provide to Ken Lunde an updated Unihan16.txt data file, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Use the updated Unihan16.txt data file to update the Unihan database to add UTC-03332 as a new `kIRG_USource` property value for U+2ECAC, based on feedback [ID20240316075059] in document L2/24-063 and Section 19 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Update the UAX #45 data file to reflect the changes to the record for UTC-03332, based on feedback [ID20240316075059] in document L2/24-063 and Section 19 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Update the UAX #45 TrueType font to map U+2ECAC to the glyph for UTC-03332, then provide the updated UAX #45 TrueType font to Michel Suignard and the Charts Working Group, based on feedback [ID20240316075059] in document L2/24-063 and Section 19 of document L2/24-067, for Unicode Version 16.0.

New:

U+2ECAC `kIRG_USource` UTC-03332

Current—*USourceData.txt*:

UTC-03332; **FutureWS**; ;75.4; ;□□木太; UTCDoc L2/22-206 35; ;8;1

Adjusted—*USourceData.txt*:

UTC-03332; **ExtI**; **U+2ECAC**; ;75.4; ;□□木太; UTCDoc L2/22-206 35; ;8;1

UAX #38 / Unihan Database Documents

The four documents in this section were reviewed and discussed by the WG, and for convenience, the following data files, which are PDF attachments, include all of the recommended Unihan database additions and changes based on them:

- *Unihan_IRGSources-changes-20240405.txt*
- *kZhuang-20240405.txt*

With the exception of *kZhuang* property value additions, the recommended Unihan database changes are also shown inline, immediately following the **Recommendations**.

20) L2/24-084: Request to confirm one unification for U+2C21C + Feedback [IRG N2636 + Feedback]

Comments: The WG previously reviewed and discussed this proposal, which resulted in the assignment of a new G-source source reference to U+2C21C, specifically **GHZR-31789.03**. The new feedback from China requested that the source reference be changed to **GIDC-F15C**. This feedback was reviewed during the IRG #62 meeting, and the IRG recommended this change. The WG also noted that this property value change necessitates a change to the normative `kIRG_GSource` property syntax that will be covered in Section 22 of this document.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the proposal to change the normative kIRG_GSource property value of U+2C21C to GIDC-F15C, based on document L2/24-084 (Chan & Tao) and Section 20 of document L2/24-067, for Unicode Version 16.0.

Action Item for Michel Suignard, CJK: Change the kIRG_GSource property value of U+2C21C to GIDC-F15C in ISO/IEC 10646's CJKSrc.txt data file, based on document L2/24-084 and Section 20 of document L2/24-067, and provide to Ken Lunde an updated Unihan16.txt data file, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Use the updated Unihan16.txt data file to update the Unihan database to change the kIRG_GSource property value for U+2C21C to GIDC-F15C, based on document L2/24-084 and Section 20 of document L2/24-067, for Unicode Version 16.0.

Current:

U+2C21C kIRG_GSource **GHRZ-31789.03**

Adjusted:

U+2C21C kIRG_GSource **GIDC-F15C**

21) L2/24-085: Request to change the radical for U+2ECAB to 130 from 74 **[IRG N2674]**

Comments: The WG reviewed and discussed this proposal, which was also reviewed and discussed during the IRG #62 meeting. The IRG recommended this *kRSUnicode* property value change that affects a single ideograph in the Extension I block, and the WG therefore also recommends this change.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the proposal to change the informative kRSUnicode property value of U+2ECAB to 130.6, based on document L2/24-085 (Chan) and Section 21 of document L2/24-067, for Unicode Version 16.0.

Action Item for Michel Suignard, CJK: Change the kRSUnicode property value of U+2ECAB to 130.6 in ISO/IEC 10646's CJKSrc.txt data file, based on document L2/24-085 and Section 21 of document L2/24-067, and provide to Ken Lunde an updated Unihan16.txt data file, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Use the updated Unihan16.txt data file to update the Unihan database to change the kRSUnicode property value of U+2ECAB to 130.6, based on document L2/24-085 and Section 21 of document L2/24-067, for Unicode Version 16.0.

Current:

U+2ECAB kRSUnicode **74.6**

Adjusted:

U+2ECAB kRSUnicode **130.6**

22) L2/24-086: Update 69 G-Source reference values [IRG N2679]

Comments: The WG reviewed and discussed this proposal, which was also reviewed and discussed during the IRG #62 meeting. The IRG recommended that the *kIRG_GSource* property values of the 69 ideographs be changed, and the WG therefore also recommends this change. The 22 GXM- and single GDM-prefixed property values will need to be zero-padded to five digits to match existing property values that use the same source prefixes.

The WG noted that the Syntax of the *kIRG_GSource* property will need to be updated to accommodate the new values that are associated with the “GIDC” prefix, which affects the following two lines of the Syntax in UAX #38:

Current:

```
| G(CE|FC|IDC23|IDC|OCD|XHZ)-\d{3}  
| GFZ-[0-9A-F]{4}
```

Adjusted:

```
| G(CE|FC|IDC23|OCD|XHZ)-\d{3}  
| G(FZ|IDC)-[0-9A-F]{4}
```

The WG also noted that the Description of the *kIRG_GSource* property’s “GIDC” prefix in UAX #38 should be changed to the following:

Supplementary Characters of the Public Security Population Information Special Font (公安人口信息专用字库补充汉字) of PRC; the value is the original PUA code point in hexadecimal

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the proposal to change the normative *kIRG_GSource* property values of 69 ideographs, based on document L2/24-086 (China) and as amended in Section 22 of document L2/24-067, for Unicode Version 16.0.

Action Item for Michel Suignard, CJK: Change the *kIRG_GSource* property values of 69 ideographs in ISO/IEC 10646’s CJKSrc.txt data file, based on document L2/24-086 and as amended in Section 22 of document L2/24-067, and provide to Ken Lunde an updated Unihan16.txt data file, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Use the updated Unihan16.txt data file to update the Unihan database to change the *kIRG_GSource* property values of 69 ideographs, based on document L2/24-086 and as amended in Section 22 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Update the Syntax and Description of the *kIRG_GSource* property in PRI #483 (Proposed Update UAX #38), based on document L2/24-086 and as amended in Section 22 of document L2/24-067, for Unicode Version 16.0.

Current:

U+2B74B	kIRG_GSource	GIDC-701
U+2B766	kIRG_GSource	GIDC-702
U+2B767	kIRG_GSource	GIDC-703
U+2B768	kIRG_GSource	GIDC-704
U+2B76A	kIRG_GSource	GIDC-705
U+2B774	kIRG_GSource	GIDC-225
U+2B785	kIRG_GSource	GIDC-706

U+2B7A0	kIRG_GSource	GIDC-707
U+2B7A2	kIRG_GSource	GIDC-708
U+2B7A7	kIRG_GSource	GIDC-709
U+2B7B7	kIRG_GSource	GIDC-710
U+2B7C7	kIRG_GSource	GIDC-711
U+2B7D1	kIRG_GSource	GIDC-712
U+2B7E0	kIRG_GSource	GIDC-713
U+2B7E2	kIRG_GSource	GIDC-714
U+2B7E4	kIRG_GSource	GIDC-606
U+2B7E5	kIRG_GSource	GIDC-715
U+2B7EB	kIRG_GSource	GIDC-716
U+2B7EC	kIRG_GSource	GIDC-717
U+2B7ED	kIRG_GSource	GIDC-623
U+2B7F3	kIRG_GSource	GIDC-718
U+2B7F4	kIRG_GSource	GIDC-719
U+2B7F5	kIRG_GSource	GIDC-720
U+2B7F6	kIRG_GSource	GIDC-721
U+2B7F7	kIRG_GSource	GIDC-722
U+2B7F8	kIRG_GSource	GIDC-723
U+2B7FB	kIRG_GSource	GIDC-724
U+2B801	kIRG_GSource	GIDC-725
U+2B802	kIRG_GSource	GIDC-726
U+2B805	kIRG_GSource	GIDC-727
U+2B80B	kIRG_GSource	GIDC-728
U+2B80F	kIRG_GSource	GIDC-729
U+2B821	kIRG_GSource	GIDC-002
U+2B82B	kIRG_GSource	GIDC-003
U+2B865	kIRG_GSource	GIDC-004
U+2B92C	kIRG_GSource	GIDC-005
U+2BA51	kIRG_GSource	GIDC-016
U+2BB59	kIRG_GSource	GIDC-008
U+2BB5B	kIRG_GSource	GIDC-010
U+2BB7D	kIRG_GSource	GIDC-011
U+2BB8C	kIRG_GSource	GIDC-012
U+2BBEE	kIRG_GSource	GIDC-015
U+2BCBB	kIRG_GSource	GIDC-055
U+2BCC1	kIRG_GSource	GIDC-030
U+2BD73	kIRG_GSource	GIDC-018
U+2BD76	kIRG_GSource	GIDC-019
U+2BD7A	kIRG_GSource	GIDC-021
U+2BD94	kIRG_GSource	GIDC-001
U+2BD9A	kIRG_GSource	GIDC-023
U+2C144	kIRG_GSource	GIDC-038
U+2C171	kIRG_GSource	GIDC-044
U+2C1F5	kIRG_GSource	GIDC-026
U+2C210	kIRG_GSource	GIDC-027
U+2C319	kIRG_GSource	GIDC-043
U+2C355	kIRG_GSource	GIDC-034
U+2C3C2	kIRG_GSource	GIDC-049
U+2C479	kIRG_GSource	GIDC-045
U+2C47A	kIRG_GSource	GIDC-046
U+2C483	kIRG_GSource	GIDC-047
U+2C72C	kIRG_GSource	GIDC-013
U+2C887	kIRG_GSource	GIDC-041
U+2C88E	kIRG_GSource	GIDC-042

U+2C9CF	kIRG_GSource	GIDC- 054
U+2CA0A	kIRG_GSource	GIDC- 031
U+2CB59	kIRG_GSource	GIDC- 050
U+2CBA2	kIRG_GSource	GIDC- 024
U+2CDFC	kIRG_GSource	GIDC- 051
U+2CE01	kIRG_GSource	GIDC- 052
U+2E00E	kIRG_GSource	GIDC- 730

Adjusted:

U+2B74B	kIRG_GSource	GXM-00434
U+2B766	kIRG_GSource	GIDC- F473
U+2B767	kIRG_GSource	GXM-00435
U+2B768	kIRG_GSource	GXM-00436
U+2B76A	kIRG_GSource	GXM-00437
U+2B774	kIRG_GSource	GIDC- E0F4
U+2B785	kIRG_GSource	GXM-00438
U+2B7A0	kIRG_GSource	GIDC- F18E
U+2B7A2	kIRG_GSource	GXM-00439
U+2B7A7	kIRG_GSource	GXM-00440
U+2B7B7	kIRG_GSource	GXM-00441
U+2B7C7	kIRG_GSource	GIDC- E51A
U+2B7D1	kIRG_GSource	GXM-00442
U+2B7E0	kIRG_GSource	GXM-00443
U+2B7E2	kIRG_GSource	GXM-00444
U+2B7E4	kIRG_GSource	GIDC- E0F0
U+2B7E5	kIRG_GSource	GXM-00445
U+2B7EB	kIRG_GSource	GXM-00446
U+2B7EC	kIRG_GSource	GXM-00447
U+2B7ED	kIRG_GSource	GIDC- E0F1
U+2B7F3	kIRG_GSource	GXM-00448
U+2B7F4	kIRG_GSource	GIDC- E57F
U+2B7F5	kIRG_GSource	GXM-00449
U+2B7F6	kIRG_GSource	GXM-00450
U+2B7F7	kIRG_GSource	GIDC- E584
U+2B7F8	kIRG_GSource	GXM-00451
U+2B7FB	kIRG_GSource	GIDC- F52B
U+2B801	kIRG_GSource	GXM-00452
U+2B802	kIRG_GSource	GXM-00453
U+2B805	kIRG_GSource	GXM-00454
U+2B80B	kIRG_GSource	GDM-00377
U+2B80F	kIRG_GSource	GXM-00455
U+2B821	kIRG_GSource	GIDC- E472
U+2B82B	kIRG_GSource	GIDC- E474
U+2B865	kIRG_GSource	GIDC- E477
U+2B92C	kIRG_GSource	GIDC- E486
U+2BA51	kIRG_GSource	GIDC- E4C0
U+2BB59	kIRG_GSource	GIDC- E497
U+2BB5B	kIRG_GSource	GIDC- E499
U+2BB7D	kIRG_GSource	GIDC- E4A0
U+2BB8C	kIRG_GSource	GIDC- E4A2
U+2BBEE	kIRG_GSource	GIDC- E4BD
U+2BCBB	kIRG_GSource	GIDC- E5CE
U+2BCC1	kIRG_GSource	GIDC- E4ED
U+2BD73	kIRG_GSource	GIDC- E4C6
U+2BD76	kIRG_GSource	GIDC- E4C7
U+2BD7A	kIRG_GSource	GIDC- E4CC

U+2BD94	kIRG_GSource	GIDC- E067
U+2BD9A	kIRG_GSource	GIDC- E4D1
U+2C144	kIRG_GSource	GIDC- E53D
U+2C171	kIRG_GSource	GIDC- E557
U+2C1F5	kIRG_GSource	GIDC- E4E2
U+2C210	kIRG_GSource	GIDC- E4E4
U+2C319	kIRG_GSource	GIDC- E552
U+2C355	kIRG_GSource	GIDC- E51F
U+2C3C2	kIRG_GSource	GIDC- E573
U+2C479	kIRG_GSource	GIDC- E56C
U+2C47A	kIRG_GSource	GIDC- E56D
U+2C483	kIRG_GSource	GIDC- E56F
U+2C72C	kIRG_GSource	GIDC- E4AC
U+2C887	kIRG_GSource	GIDC- E54D
U+2C88E	kIRG_GSource	GIDC- E54F
U+2C9CF	kIRG_GSource	GIDC- E5BE
U+2CA0A	kIRG_GSource	GIDC- E4F3
U+2CB59	kIRG_GSource	GIDC- E591
U+2CBA2	kIRG_GSource	GIDC- E4DA
U+2CDFC	kIRG_GSource	GIDC- E59D
U+2CE01	kIRG_GSource	GIDC- E59E
U+2E00E	kIRG_GSource	GIDC- E5F5

23) L2/24-089: Proposal to add kZhuang to Unihan [IRG N2677]

Comments: The WG reviewed and discussed this proposal, which is to add a new provisional *kZhuang* property to the Unihan database that provides Zhuang readings for nearly 2,500 ideographs in the Extensions F through H blocks. The WG noted that we added the provisional *kZhuangNumeric* property in Unicode Version 15.1, so there is precedent for adding the equivalent reading property.

Recommendations: The WG recommends to the UTC the following disposition ⇨

Consensus: Accept the new provisional Unihan database property, *kZhuang*, based on document L2/24-089 (Knightley) and Section 23 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Document the *kZhuang* property in PRI #483 (Proposed Update UAX #38), based on document L2/24-089 and Section 23 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Populate the *kZhuang* property in the Unihan database with the property values as provided in document L2/24-089, based on Section 23 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the author of document L2/24-089 the CJK & Unihan Working Group comments in Section 23 of document L2/24-067.

UAX #45 / U-Source Documents

The seven documents in this section were reviewed and discussed by the WG, and for convenience, the following data file, which is a PDF attachment, includes the recommended UAX #45 *USourceData.txt* data file additions based on it:

- *USource-additions-20240405.txt*

The recommended UAX #45 *USourceData.txt* data file additions are also shown inline, immediately following the **Recommendations**.

24) L2/24-041: Proposal to add 10 characters used in Putian-Xianyou Puxianxi Opera to UAX #45

Comments: The WG reviewed and discussed this proposal, and agreed that adding the ideographs to UAX #45 is reasonable, with an assigned status value of **FutureWS**. The ideographs, whose representative glyphs are shown below, can therefore be considered as candidates for inclusion in the UTC's submission for IRG Working Set 2024:

嚙 嘜 揸 砑 嘍 嚙 嚙 屍 嚙 倭

Recommendations: The WG recommends to the UTC the following disposition ⇨

Consensus: Accept 10 new U-source ideographs as UTC-03399 through UTC-03408 with a UAX #45 status value of FutureWS, based on document L2/24-041 (Jiang) and Section 24 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Add 10 new records to the UAX #45 *USourceData.txt* data file and their representative glyphs to *USourceGlyphs.pdf*, based on document L2/24-041 and Section 24 of document L2/24-067, for Unicode Version 16.0.

New:

```
UTC-03399;FutureWS;;30.15;;𠄎;UTCDoc L2/24-041 1;PX Min /iã42/;18;1
UTC-03400;FutureWS;;30.13;;𠄏;UTCDoc L2/24-041 2;PX Min /hui453/;16;3
UTC-03401;FutureWS;;64.8;;𠄐;UTCDoc L2/24-041 3;PX Min /ŋia24/;11;3
UTC-03402;FutureWS;;118.5;;𠄑;UTCDoc L2/24-041 4;PX Min /tu453/;11;1
UTC-03403;FutureWS;;30.15;;𠄒;UTCDoc L2/24-041 5;PX Min /hui533/;18;2
UTC-03404;FutureWS;;30.13;;𠄓;UTCDoc L2/24-041 6;PX Min /tshãh11/;16;2
UTC-03405;FutureWS;;30.17;;𠄔;UTCDoc L2/24-041 7;kSimplifiedVariant
U+20DE9;20;1
UTC-03406;FutureWS;;44.5;;𠄕;UTCDoc L2/24-041 8;PX Min /khɔ533/;8;1
UTC-03407;FutureWS;;30.11;;𠄖;UTCDoc L2/24-041 9;PX Min /khã42/;14;1
UTC-03408;FutureWS;;9.9;;𠄗;UTCDoc L2/24-041 10;;11;3
```

25) L2/24-047: Proposal to add 34 characters related to fish and fish product in Chinese to UAX #45

Comments: The WG reviewed and discussed this proposal, and agreed that adding the ideographs to UAX #45 is reasonable, with an assigned status value of **FutureWS**. The ideographs, whose representative glyphs are shown below, can therefore be considered as candidates for inclusion in the UTC’s submission for IRG Working Set 2024:

𩚑 𩚒 𩚓 𩚔 𩚕 𩚖 𩚗 𩚘 𩚙 𩚚 𩚛 𩚜 𩚝 𩚞 𩚟 𩚠 𩚡 𩚢 𩚣 𩚤 𩚥 𩚦 𩚧 𩚨 𩚩 𩚪 𩚫 𩚬 𩚭 𩚮 𩚯 𩚰 𩚱 𩚲 𩚳 𩚴 𩚵 𩚶 𩚷 𩚸 𩚹 𩚺 𩚻 𩚼 𩚽 𩚾 𩚿

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept 34 new U-source ideographs as UTC-03409 through UTC-03442 with a UAX #45 status value of FutureWS, based on document L2/24-047 (Chan) and Section 25 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Add 34 new records to the UAX #45 USourceData.txt data file and their representative glyphs to USourceGlyphs.pdf, based on document L2/24-047 and Section 25 of document L2/24-067, for Unicode Version 16.0.

New:

- UTC-03409; FutureWS; ;195'.4; ; 𩚑 鱼友; UTCDoc L2/24-047 1; kTraditionalVariant U+29D7C: kCantonese jau5;12;1
- UTC-03410; FutureWS; ;195'.4; ; 𩚒 鱼今; UTCDoc L2/24-047 2; kTraditionalVariant U+4C3C: kMandarin qín: kCantonese cam4;12;3
- UTC-03411; FutureWS; ;195'.4; ; 𩚓 鱼公; UTCDoc L2/24-047 3; kTraditionalVariant U+4C38: kMandarin gōng: kCantonese gung1;12;3
- UTC-03412; FutureWS; ;195'.4; ; 𩚔 鱼勿; UTCDoc L2/24-047 4; kTraditionalVariant U+9B69: kMandarin mò: kCantonese mut6;12;3
- UTC-03413; FutureWS; ;195'.6; ; 𩚕 鱼产; UTCDoc L2/24-047 5; kTraditionalVariant U+29ECF;14;4
- UTC-03414; FutureWS; ;195'.4; ; 𩚖 鱼历; UTCDoc L2/24-047 6; kTraditionalVariant U+29F4F: kMandarin lì: Min Dong lik⁴;12;1
- UTC-03415; FutureWS; ;195'.2; ; 𩚗 鱼力; UTCDoc L2/24-047 7; kTraditionalVariant U+29D53: kMandarin lèi;10;5
- UTC-03416; FutureWS; ;195'.4; ; 𩚘 鱼户; UTCDoc L2/24-047 8; kTraditionalVariant U+9B72: kSemanticVariant U+9C78 U+9C88: kMandarin lú: kCantonese lou4;12;3
- UTC-03417; FutureWS; ;195'.9; ; 𩚙 鱼若; UTCDoc L2/24-047 9; kTraditionalVariant U+9C19: kMandarin ruò: kCantonese joek6: Wu n̩iæ?^{2 2} n̩ia?² n̩ie?^{1 2};17;2
- UTC-03418; FutureWS; ;195'.13; ; 𩚚 鱼辟; UTCDoc L2/24-047 10; kTraditionalVariant U+29F0E;21;5
- UTC-03419; FutureWS; ;195'.8; ; 𩚛 鱼定; UTCDoc L2/24-047 11; kTraditionalVariant U+29E0E: kMandarin dīng;16;4
- UTC-03420; FutureWS; ;195'.8; ; 𩚜 鱼卖; UTCDoc L2/24-047 12; kTraditionalVariant U+29F46: kMandarin xù;16;1

UTC-03421;FutureWS;;195'.3;;𩺰;UTCDoc L2/24-047 13;kTraditionalVariant
 U+4C35;kMandarin zī;kCantonese zil;11;5
 UTC-03422;FutureWS;;195'.6;;𩺱;UTCDoc L2/24-047 14;;14;1
 UTC-03423;FutureWS;;195'.6;;𩺲;UTCDoc L2/24-047 15;kTraditionalVariant
 U+29DB8;14;2
 UTC-03424;FutureWS;;195'.7;;𩺳;UTCDoc L2/24-047 16;kTraditionalVariant
 U+4C55;kMandarin móu;kCantonese mau4;15;3
 UTC-03425;FutureWS;;195'.6;;𩺴;UTCDoc L2/24-047 17;kTraditionalVariant
 U+2CD4F;kMandarin shè;14;3
 UTC-03426;FutureWS;;195'.10;;𩺵;UTCDoc L2/24-047 18;kTraditionalVariant
 U+2B66E;18;3
 UTC-03427;FutureWS;;195'.9;;𩺶;UTCDoc L2/24-047 19;kTraditionalVariant WS2021-
 04607;kMandarin hòu;17;1
 UTC-03428;FutureWS;;195'.5;;𩺷;UTCDoc L2/24-047 20;kTraditionalVariant
 U+9B94;kMandarin jù;kCantonese geoi6;13;1
 UTC-03429;FutureWS;;195'.5;;𩺸;UTCDoc L2/24-047 21;kTraditionalVariant
 U+29D97;13;3
 UTC-03430;FutureWS;;195'.9;;𩺹;UTCDoc L2/24-047 22;kTraditionalVariant
 U+29E8C;17;1
 UTC-03431;FutureWS;;195'.11;;𩺺;UTCDoc L2/24-047 23;kTraditionalVariant
 U+2B672;19;2
 UTC-03432;FutureWS;;195'.9;;𩺻;UTCDoc L2/24-047 24;;17;2
 UTC-03433;FutureWS;;195'.7;;𩺼;UTCDoc L2/24-047 25;kTraditionalVariant
 U+4C54;kMandarin tiáo;kCantonese tiu4;15;3
 UTC-03434;FutureWS;;195'.5;;𩺽;UTCDoc L2/24-047 26;kTraditionalVariant
 U+9C71;kMandarin lì;kCantonese lai6;kVietnamese lịch;13;1
 UTC-03435;FutureWS;;195'.10;;𩺾;UTCDoc L2/24-047 27;kTraditionalVariant
 U+2B681;18;1
 UTC-03436;FutureWS;;195'.8;;𩺿;UTCDoc L2/24-047 28;kTraditionalVariant
 U+29F3A;16;2
 UTC-03437;FutureWS;;195'.4;;𩻀;UTCDoc L2/24-047 29;kTraditionalVariant
 U+9B70;kMandarin wén;kCantonese man4;12;4
 UTC-03438;FutureWS;;195'.12;;𩻁;UTCDoc L2/24-047 30;kTraditionalVariant
 U+29EF8;kCantonese dai6;20;5
 UTC-03439;FutureWS;;195'.9;;𩻂;UTCDoc L2/24-047 31;kTraditionalVariant
 U+9C0E;kMandarin jiǎn;kCantonese gin2;17;5
 UTC-03440;FutureWS;;195'.7;;𩻃;UTCDoc L2/24-047 32;kTraditionalVariant
 U+29DDA;kMandarin bó;15;1
 UTC-03441;FutureWS;;195'.10;;𩻄;UTCDoc L2/24-047 33;kTraditionalVariant
 U+29E76;kMandarin táng;18;4
 UTC-03442;FutureWS;;195'.4;;𩻅;UTCDoc L2/24-047 34;kTraditionalVariant
 U+9C38;kMandarin qū;kCantonese keoi1;12;1

26) L2/24-056: Proposal to add three characters used for Cantonese education to UAX #45

Comments: The WG reviewed and discussed this proposal, and agreed that adding the ideographs to UAX #45 is reasonable, with an assigned status value of **FutureWS**. The WG noted that the first ideograph is the simplified form of an ideograph whose *kCantonese* property value was discussed in Section 07 of this document. The ideographs, whose representative glyphs are shown below, can therefore be considered as candidates for inclusion in the UTC's submission for IRG Working Set 2024:

𨮑𨮑𨮑

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept three new U-source ideographs as UTC-03443 through UTC-03445 with a UAX #45 status value of FutureWS, based on document L2/24-056 (Chan & Liang) and Section 26 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Add three new records to the UAX #45 USourceData.txt data file and their representative glyphs to USourceGlyphs.pdf, based on document L2/24-056 and Section 26 of document L2/24-067, for Unicode Version 16.0.

New:

UTC-03443;FutureWS;;159'.4;;𨮑𨮑𨮑𨮑;UTCDoc L2/24-056 1;kTraditionalVariant
U+282CD:kCantonese wen1;8;5

UTC-03444;FutureWS;;130.6;;𨮑𨮑𨮑𨮑;UTCDoc L2/24-056 2;kCantonese dyut1;10;1

UTC-03445;FutureWS;;30.10;;𨮑𨮑𨮑𨮑;UTCDoc L2/24-056 3;kCantonese lyun1;13;4

27) L2/24-075: Proposal to add one character used for Cantonese particle to UAX #45

Comments: The WG reviewed and discussed this proposal, and agreed that adding the ideograph to UAX #45 is reasonable, with an assigned status value of **FutureWS**. The ideograph, whose representative glyph is shown below, can therefore be considered as a candidate for inclusion in the UTC's submission for IRG Working Set 2024:

𨮑𨮑

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept one new U-source ideograph as UTC-03446 with a UAX #45 status value of FutureWS, based on document L2/24-075 (Ng) and Section 27 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Add one new record to the UAX #45 USourceData.txt data file and its representative glyph to USourceGlyphs.pdf, based on document L2/24-075 and Section 27 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the author of document L2/24-075 the CJK & Unihan Working Group comments in Section 27 of document L2/24-067.

New:

UTC-03446;FutureWS;;30.10;;𨮑𨮑𨮑𨮑;UTCDoc L2/24-075;kTraditionalVariant
U+21075:kCantonese gan2;13;2

28) L2/24-076: Proposal to add three characters to UAX #45

Comments: The WG reviewed and discussed this proposal, and agreed that adding the ideographs to UAX #45 is reasonable, with an assigned status value of **FutureWS**. The WG noted that adding [KC05766](#), which is mentioned on the second page of the proposal, to the metadata of the third ideograph would be helpful. The ideographs, whose representative glyphs are shown below, can therefore be considered as candidates for inclusion in the UTC's submission for IRG Working Set 2024:

馬占 王品 車丘

Recommendations: The WG recommends to the UTC the following disposition ⇨

Consensus: Accept three new U-source ideographs as UTC-03447 through UTC-03449 with a UAX #45 status value of FutureWS, based on document L2/24-076 (Ren) and as amended in Section 28 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Add three new records to the UAX #45 USourceData.txt data file and their representative glyphs to USourceGlyphs.pdf, based on document L2/24-076 and as amended in Section 28 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the author of document L2/24-076 the CJK & Unihan Working Group comments in Section 28 of document L2/24-067.

New:

UTC-03447;FutureWS;;187.5;;馬占;UTCDoc L2/24-076 1;kMandarin zhàn;15;2
UTC-03448;FutureWS;;96.9;;王品;UTCDoc L2/24-076 2;;13;2
UTC-03449;FutureWS;;159.5;;車丘;UTCDoc L2/24-076 3;U+9A45:KC05766;12;3

29) L2/24-092: Proposal to add 28 ideographs used in All Men Are Brothers and Classic of Mountains and Rivers to UAX#45

Comments: The WG reviewed and discussed this proposal, and agreed that adding the ideographs to UAX #45 is reasonable, with an assigned status value of **FutureWS**. The WG noted that the 14th ideograph in this proposal, 𩺰执鱼, is in IRG Working Set 2021 as serial number [04667](#) (aka GKJ-00283), and should therefore be excluded. The remaining 27 ideographs, whose representative glyphs are shown below, can therefore be considered as candidates for inclusion in the UTC's submission for IRG Working Set 2024:

蚩梢獯𧯂𧯃𧯄𧯅𧯆𧯇𧯈𧯉𧯊𧯋𧯌𧯍𧯎𧯏𧯐𧯑𧯒𧯓𧯔𧯕
𧯖𧯗𧯘𧯙𧯚𧯛𧯜𧯝𧯞𧯟𧯠𧯡𧯢𧯣𧯤

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept 27 new U-source ideographs as UTC-03450 through UTC-03476 with a UAX #45 status value of FutureWS, based on document L2/24-092 (Wang & He) and as amended in Section 29 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Add 27 new records to the UAX #45 USourceData.txt data file and their representative glyphs to USourceGlyphs.pdf, based on document L2/24-092 and as amended in Section 29 of document L2/24-067, for Unicode Version 16.0.

New:

UTC-03450;FutureWS;;;46.6;;;𪛗山虫;UTCDoc L2/24-092 1;kMandarin chī:kSemanticVariant U+86A9;9;2
UTC-03451;FutureWS;;;75.10;;;𪛘木宵;UTCDoc L2/24-092 2;kMandarin lǔ:kSemanticVariant U+6AD3;14;4
UTC-03452;FutureWS;;;94.12;;;𪛙豸鬣;UTCDoc L2/24-092 3;kMandarin xié:kTraditionalVariant U+248BA;15;1
UTC-03453;FutureWS;;;115.4;;;𪛚禾疋;UTCDoc L2/24-092 4;kMandarin zǒu:kSemanticVariant U+27E86 U+8D70;9;2
UTC-03454;FutureWS;;;118.9;;;𪛛竹俛;UTCDoc L2/24-092 5;kMandarin jué:kSemanticVariant U+855D;15;3
UTC-03455;FutureWS;;;118.9;;;𪛜竹绝;UTCDoc L2/24-092 6;kMandarin jué:kTraditionalVariant U+25CF5:kSemanticVariant U+2B235;15;5
UTC-03456;FutureWS;;;119.9;;;𪛝米费;UTCDoc L2/24-092 7;kMandarin pì:kTraditionalVariant U+42A7:kSemanticVariant U+5C41;15;5
UTC-03457;FutureWS;;;120'.8;;;𪛞彡昔;UTCDoc L2/24-092 8;kMandarin zuó:kTraditionalVariant U+2604E;11;1
UTC-03458;FutureWS;;;123.5;;;𪛟羊东;UTCDoc L2/24-092 9;kMandarin dōng:kTraditionalVariant U+4376;11;1
UTC-03459;FutureWS;;;142.5;;;𪛠虫鸟;UTCDoc L2/24-092 10;kMandarin tóng:kTraditionalVariant U+4CCB;11;3
UTC-03460;FutureWS;;;159'.9;;;𪛡车扁;UTCDoc L2/24-092 11;kMandarin biàn:kTraditionalVariant U+4862;13;4
UTC-03461;FutureWS;;;169'.8;;;𪛢门豕;UTCDoc L2/24-092 12;kMandarin dū:kTraditionalVariant U+49A0;11;1
UTC-03462;FutureWS;;;187'.9;;;𪛣马券;UTCDoc L2/24-092 13;kMandarin téng:kTraditionalVariant U+9A63:kSemanticVariant U+9A30;12;4
UTC-03463;FutureWS;;;195'.10;;;𪛤鱼尊;UTCDoc L2/24-092 15;kMandarin bū:kTraditionalVariant U+29E72;13;1
UTC-03464;FutureWS;;;195'.11;;;𪛥鱼巢;UTCDoc L2/24-092 16;kMandarin zǎo:kTraditionalVariant U+4C83;19;5
UTC-03465;FutureWS;;;196'.3;;;𪛦鸟大;UTCDoc L2/24-092 17;kMandarin dài:kTraditionalVariant U+4CA6;8;1
UTC-03466;FutureWS;;;196'.5;;;𪛧乐鸟;UTCDoc L2/24-092 18;kMandarin luò:kTraditionalVariant U+2A1F1;10;3
UTC-03467;FutureWS;;;196'.5;;;𪛨丘鸟;UTCDoc L2/24-092 19;kMandarin chī:kTraditionalVariant U+29FE8:kSemanticVariant U+9D1F;10;3
UTC-03468;FutureWS;;;196'.6;;;𪛩次鸟;UTCDoc L2/24-092 20;kMandarin cì:kTraditionalVariant U+4CD0;11;4
UTC-03469;FutureWS;;;196'.8;;;𪛪鸟一宛;UTCDoc L2/24-092 21;kMandarin wǎn:kTraditionalVariant U+2A0A7;12;4
UTC-03470;FutureWS;;;196'.10;;;𪛫贲鸟;UTCDoc L2/24-092 22;kMandarin fén:kTraditionalVariant U+4D05;15;1
UTC-03471;FutureWS;;;196'.10;;;𪛬般鸟;UTCDoc L2/24-092 23;kMandarin pán:kTraditionalVariant U+2A100;15;3

UTC-03472;FutureWS;;196'.11;;𪛗;UTCDoc L2/24-092 24;kMandarin
xí:kTraditionalVariant U+2A136;16;5
UTC-03473;FutureWS;;196'.13;;𪛘;UTCDoc L2/24-092 25;kMandarin
chù:kTraditionalVariant U+2A1C6;18;2
UTC-03474;FutureWS;;196'.13;;𪛙;UTCDoc L2/24-092 26;kMandarin
dǎn:kTraditionalVariant U+2A1BB;18;3
UTC-03475;FutureWS;;196'.13;;𪛚;UTCDoc L2/24-092 27;kMandarin
dié:kTraditionalVariant U+4D11;18;5
UTC-03476;FutureWS;;196'.16;;𪛛;UTCDoc L2/24-092 28;kMandarin
hōng:kTraditionalVariant U+2A218;21;1

30) L2/24-095: Proposal to add one character used for Chinese ancient books arrangement to UAX #45

Comments: The WG reviewed and discussed this proposal, and agreed that adding the ideographs to UAX #45 is reasonable, with an assigned status value of **FutureWS**. The WG decided that adding [KC07952](#), which is mentioned on the first page of the proposal, to its metadata would be helpful. The ideograph, whose representative glyph is shown below, can therefore be considered as a candidate for inclusion in the UTC's submission for IRG Working Set 2024:

𪛛

Recommendations: The WG recommends to the UTC the following disposition ⇨

Consensus: Accept one new U-source ideograph as UTC-03477 with a UAX #45 status value of FutureWS, based on document L2/24-095 (Zhao) and as amended in Section 30 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Add one new record to the UAX #45 USourceData.txt data file and its representative glyph to USourceGlyphs.pdf, based on document L2/24-095 and as amended in Section 30 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the author of document L2/24-095 the CJK & Unihan Working Group comments in Section 30 of document L2/24-067.

New:

UTC-03477;FutureWS;;27.9;;𪛛;UTCDoc L2/24-095;kMandarin bì:kSemanticVariant
U+6E62:KC07952;11;1

31) L2/24-099: Proposal to add 10 ideographs used in the Secret History of the Mongols

Comments: The WG reviewed and discussed this proposal, and agreed that adding the ideographs to UAX #45 is reasonable, with an assigned status value of **FutureWS**. The ideographs, whose representative glyphs are shown below, can therefore be considered as candidates for inclusion in the UTC's submission for IRG Working Set 2024:

𪛜 𪛝 𪛞 𪛟 𪛠 𪛡 𪛢 𪛣 𪛤 𪛥

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept 10 new U-source ideographs as UTC-03478 through UTC-03487 with a UAX #45 status value of FutureWS, based on document L2/24-099 (Yehenara & Sim) and Section 31 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Add 10 new records to the UAX #45 USourceData.txt data file and their representative glyphs to USourceGlyphs.pdf, based on document L2/24-099 and Section 31 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the author of document L2/24-099 the CJK & Unihan Working Group comments in Section 31 of document L2/24-067.

New:

UTC-03478;FutureWS;;109.16;;𠃉目𠃉𠃉王𠃉𠃉;UTCDoc L2/24-099 1;U+77D1;21;1
UTC-03479;FutureWS;;157.14;;𠃉𠃉𠃉𠃉;UTCDoc L2/24-099 2;U+8E94;21;1
UTC-03480;FutureWS;;169.3;;𠃉門𠃉𠃉;UTCDoc L2/24-099 3;;11;1
UTC-03481;FutureWS;;187.19;;𠃉馬羅;UTCDoc L2/24-099 4;;28;2
UTC-03482;FutureWS;;123.7;;𠃉羊𠃉;UTCDoc L2/24-099 5;;12;2
UTC-03483;FutureWS;;196.13;;𠃉窟鳥;UTCDoc L2/24-099 6;;24;3
UTC-03484;FutureWS;;196.7;;𠃉克鳥;UTCDoc L2/24-099 7;;18;1
UTC-03485;FutureWS;;208.9;;𠃉鼠革;UTCDoc L2/24-099 8;;22;1
UTC-03486;FutureWS;;208.9;;𠃉巢草;UTCDoc L2/24-099 9;;19;1
UTC-03487;FutureWS;;208.16;;𠃉巢𠃉𠃉丁𠃉𠃉𠃉了𠃉𠃉二𠃉;UTCDoc L2/24-099 10;U+2A58C;26;1

Other Documents

The 10 documents in this section were reviewed and discussed by the WG, and for convenience, the following data file, which is a PDF attachment, includes the recommended *EquivalentUnifiedIdeograph.txt* data file additions based on it:

- *EquivalentUnifiedIdeograph-20240405.txt*

The recommended *EquivalentUnifiedIdeograph.txt* data file additions are also shown inline, immediately following the **Recommendations**.

32) L2/24-044: Correcting the proposed name of new character U+31E4 in Unicode 16.0

Comments: The WG reviewed and discussed this proposal, and agreed that changing the name of U+31E4 CJK STROKE HZXG, specifically to remove the “Z,” is appropriate. For background, this new character in the [CJK Strokes](#) block was accepted for Unicode Version 16.0 during the UTC #177 meeting per [Consensus 177-C13](#). As stated in the proposal, this character name change affects several data files. This character name change should also be included in ballot comments to WG2. The proposal author already provided to the WG chair the updated SVG file for Appendix F of the Core Specification. In addition to the data files that were cited in the proposal, the WG noted that two commented-out lines should be appended to the UCD’s

EquivalentUnifiedIdeograph.txt data file (these two characters do not have an equivalent unified ideograph, but should be mentioned in the commented-out section at the end of the data file).

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the proposal to change the name of U+31E4, from CJK STROKE HZXG to CJK STROKE HXG, based on document L2/24-044 (Koo) and Section 32 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Whistler, UTC: Update the Pipeline to change the character name of U+31E4 to CJK STROKE HXG, based on document L2/24-044 and Section 32 of document L2/24-067, for Unicode Version 16.0.

Action Item for Debbie Anderson, UTC: Add to WG2 ballot comments the character name change of U+31E4, from CJK STROKE HZXG to CJK STROKE HXG, based on document L2/24-044 and as amended in Section 32 of document L2/24-067, for Unicode Version 16.0.

Action Item for Markus Scherer, PAG: Update the UCD's *EquivalentUnifiedIdeograph.txt* data file to append two lines for U+31E4 and U+31E5, based on document L2/24-044 and Section 32 of document L2/24-067 and its attached *EquivalentUnifiedIdeograph-20240405.txt* data file, for Unicode Version 16.0.

Action Item for Ken Whistler, EDC: Update *NamesList.txt* to change the character name of U+31E4 to CJK STROKE HXG, based on document L2/24-044 and Section 32 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, EDC: Add to Editorial Working Group GitHub Issue #65 the updated SVG file for U+31E4, based on document L2/24-044 and as amended in Section 32 of document L2/24-067, for Unicode Version 16.0.

Action Item for Ken Lunde, CJK: Convey to the author of document L2/24-044 the CJK & Unihan Working Group comments in Section 32 of document L2/24-067.

New—*EquivalentUnifiedIdeograph.txt*:

```
# 31E4; CJK STROKE HXG  
# 31E5; CJK STROKE SZP
```

33) L2/24-053R: US/Unicode Activity Report for IRG #62 Meeting [IRG N2667R]

Comments: This document is the UTC's activity report as delivered during the IRG #62 meeting in March of 2024, and is therefore an FYI document for the UTC. This document was revised on 2024-03-17 to add the fifth bullet that provided an estimate of the size of the UTC's submission for IRG Working Set 2024.

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

34) L2/24-082: IRG Meeting #62 Recommendations and Action Items [IRG N2670]

Comments: While this is mainly an FYI document for the UTC, because all of the necessary actions have been specified under documents that are referenced elsewhere in this document, it is important to point out the future IRG meeting schedule, the IRG Working Set 2021 schedule, and proposals that are likely to reach the UTC during the UTC #181 meeting in November, all of which impact UTC experts:

- Per Recommendation IRG M62.01, the next IRG meeting, **IRG #63 (2024-10-21 through 2024-10-25)** will be hosted by ROK in Seoul (hybrid); **IRG #64 (2025-03-17 through 2025-03-21)** will be virtual (Zoom) and hosted by the Unicode Consortium; **IRG #65 (2025-10-13 through 2025-10-17)** will be hosted by China in Guangzhou (hybrid); and **IRG #66 (2026-03-16 through 2026-03-20)** will be virtual (Zoom) and hosted by the Unicode Consortium. None of the dates have changed.
- Per Recommendation IRG M62.05, the IRG Convenor is planning to submit **IRG Working Set 2021 Version 7.0**, which will include over 4,000 ideographs, to WG2 by **2023-05-31**, which effectively confirms that the corresponding new CJK Unified Ideographs Extension J block could be reasonably targeted for Unicode Version 17.0 (2025).
- Per Recommendation IRG M62.06, and as proposed in document [IRG N2643R2](#), ROK's 306-ideograph horizontal extension is likely to be accepted by the IRG during the IRG #63 meeting in October, meaning that it may be a topic during the UTC #181 meeting in November.
- Per Recommendation IRG M62.07, and as proposed in document [IRG N2676](#), U+5CC0 𪗀 is likely to be disunified by the IRG during the IRG #63 meeting in October, meaning that it may be a topic during the UTC #181 meeting in November.
- Per Recommendation IRG M62.12, and as proposed in document [IRG N2656](#), China is likely to accept or adjust the proposed *kIRG_GSource* property source reference changes during the IRG #63 meeting in October, meaning that it may be a topic during the UTC #181 meeting in November.

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

35) L2/24-083: Editorial Report on Miscellaneous Issues [IRG N2671]

Comments: This is an FYI document for the UTC, because all of the necessary actions in Sections 1, 6, and 7 are covered by Sections 20, 22, and 21 of this document, respectively. The last bullet in Section 3 states that SAT may provide an updated S-source representative glyph for U+2D156, but the soonest that this would happen is during the IRG #63 meeting in October. Shown below is the code chart excerpt for U+2D156 as it appears in Unicode Version 16.0 Alpha, which reflects the adjusted G-source representative glyph:

<p>2D156</p> <p>𠄎 22.3</p> <p>𠄎 96.2</p>	 GKJ-00102	 SAT-60036
--	--	---

For reference, the code chart excerpt for the same ideograph as it appears in Unicode Versions 14.0 through 15.1 is shown below (the G-source source reference and its representative glyph were added in Version 14.0):

<p>2D156</p> <p>𠄎 22.3</p>	 GKJ-00102	 SAT-60036
----------------------------	--	---

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

36) L2/24-057: Proposal of the UNICODE AUTO SPACING

Comments: The WG reviewed and discussed this proposal, which is a revised version of document L2/23-283 that was included in Section 20 of document L2/24-012 (*CJK & Unihan Group Recommendations for UTC #178 Meeting*) and discussed during the UTC #178 meeting in January. As noted in Section 20 of L2/24-012, this revised proposal is still a preliminary proposal, because the data file is still being refined, and various issues are still being discussed on GitHub. One of the proposal authors, Koji Ishii (石井宏治), was present during the meeting, and led the discussion. Based on feedback during the UTC #178 meeting, *East_Asian_Width* property values, which are normative from Unicode Version 13.0 (2020), are now being used to derive the values for this proposed new property, which may result in recommended changes to some *East_Asian_Width* property values.

The WG noted that the acceptable use of explicit “space” or virtual “space” has the potential to cause interoperability issues. The WG further noted that Hangul syllables and related characters that behave like Hiragana and Katakana, along with Yi syllables, should probably be assigned “W,” not “O.” Hangul is a bit tricky due to the use of spaces in modern Korean text. The image below, which is an excerpt from *大学彝语文：彝文* (洛边木果, 四川民族出版社, 2000), shows spacing between Yi syllables and digits:

ຈົນຖານໄດ້ຮັບການຊີ້ນຳ, ຈົນຖານໄດ້ຮັບການຊີ້ນຳ, ຈົນຖານໄດ້ຮັບການຊີ້ນຳ, ຈົນຖານໄດ້ຮັບການຊີ້ນຳ, ຈົນຖານໄດ້ຮັບການຊີ້ນຳ, ຈົນຖານໄດ້ຮັບການຊີ້ນຳ. ຈົນຖານໄດ້ຮັບການຊີ້ນຳ
 “ຊີ້ນຳຊີ້ນຳ” ຈົນຖານໄດ້ຮັບການຊີ້ນຳ ຈົນຖານໄດ້ຮັບການຊີ້ນຳ: ຈົນຖານໄດ້ຮັບການຊີ້ນຳ
 5—5 ຈົນຖານ“55 ຈົນຖານ” ຈົນຖານ; ຈົນຖານໄດ້ຮັບການຊີ້ນຳ 4—4 ຈົນ
 ຖານ“44 ຈົນຖານ” ຈົນຖານ; ຈົນຖານໄດ້ຮັບການຊີ້ນຳ 3—3 ຈົນຖານ“33 ຈົນຖານ”
 ຈົນຖານ, ຈົນຖານໄດ້ຮັບການຊີ້ນຳ 2 ຈົນຖານ 1 ຈົນຖານ“21 ຈົນຖານ” ຈົນຖານ.

About the three proposed property values—**W**, **N**, and **O**—the WG noted that it would be helpful to provide short names for the purpose of discussing them. It is clear that “W” means “Wide.” Does “N” mean “Narrow” or “Neutral”? Does “O” mean “Other”?

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

37) L2/24-059: Proposal to change the East_Asian_Width property of the Yijing symbols

L2/24-072: Proposal to change the East_Asian_Width property of East Asian numeric characters

Comments: The WG reviewed and discussed both of these documents, which collectively propose to change the normative *East_Asian_Width* property values of the following 188 characters from **N** (*Neutral*) to **W** (*East Asian Wide*):

L2/24-059: U+2630..U+2637, U+268A..U+268F, U+4DC0..U+4DFF, U+1D300..U+1D356 (165)

L2/24-072: U+1D360..U+1D376 (23)

The WG noted that all 188 characters have a *Vertical_Orientation* property value of **U** (*Upright*), and when implemented in fonts, their glyphs are likely to have full-width metrics. The WG therefore recommends that their *East_Asian_Width* property values be changed from **N** to **W**, and that a new PRI be issued. There has been some discussion about making some adjustments to the text in [UAX #11](#) (*East Asian Width*), which should be included in the same PRI.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the proposal to change the *East_Asian_Width* property values of 188 characters from N to W, based on documents L2/24-059 and L2/24-072 (Bertholet) and Section 37 of document L2/24-067, for Unicode Version 16.0.

Action Item for Markus Scherer, PAG: Update the UCD’s *EastAsianWidth.txt* data file to change the *East_Asian_Width* property values of 188 characters from N to W, based on documents L2/24-059 and L2/24-072 and Section 37 of document L2/24-067.

Action Item for Ken Lunde, CJK: Prepare a new PRI for a Proposed Update of UAX #11 that records the East_Asian_Width property value changes for 188 characters from N to W, based on documents L2/24-059 and L2/24-072 and Section 37 of document L2/24-067.

Action Item for Rick McGowan, UTC: Post a new PRI for the Proposed Update of UAX #11 with a closing date of 2024-07-02.

Action Item for Ken Lunde, CJK: Convey to the author of documents L2/24-059 and L2/24-072 the CJK & Unihan Working Group comments in Section 37 of document L2/24-067.

38) L2/24-071R: Proposal to encode three stable extended Suzhou Numeral-like letters for Cantonese Music

Comments: The WG reviewed and discussed the original version of this document, which is a proposal to encode three characters that are used for Cantonese music notation. The WG agreed that there was sufficient evidence to encode these characters, and to provisionally assign code points for them in the [Ideographic Symbols and Punctuation](#) block, which looked reasonable. However, the WG felt that it was inappropriate to include **SUZHOU NUMERAL** as part of the character names, primarily because they are not numerals. A revised version of this document changed the character names to **YANGQIN SIGN SLOW TWO**, **YANGQIN SIGN SLOW THREE**, and **YANGQIN SIGN SLOW FOUR**, which are both appropriate and accurate.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Consensus: Accept the proposal to provisionally assign U+16FF4 YANGQIN SIGN SLOW TWO, U+16FF5 YANGQIN SIGN SLOW THREE, and U+16FF6 YANGQIN SIGN SLOW FOUR in the Ideographic Symbols and Punctuation block, based on document L2/24-071R (Chan) and Section 38 of document L2/24-067.

Action Item for Ken Whistler, UTC: Update the Pipeline to add the three provisionally assigned characters U+16FF4 YANGQIN SIGN SLOW TWO, U+16FF5 YANGQIN SIGN SLOW THREE, and U+16FF6 YANGQIN SIGN SLOW FOUR, based on document L2/24-071R and Section 38 of document L2/24-067.

Action Item for Ken Lunde, CJK: Provide a TrueType font for U+16FF4 YANGQIN SIGN SLOW TWO, U+16FF5 YANGQIN SIGN SLOW THREE, and U+16FF6 YANGQIN SIGN SLOW FOUR to Michel Suignard and the Charts Working Group, based on document L2/24-071R and Section 38 of document L2/24-067.

Action Item for Ken Lunde, CJK: Prepare text for the Core Specification that describes U+16FF4 YANGQIN SIGN SLOW TWO, U+16FF5 YANGQIN SIGN SLOW THREE, and U+16FF6 YANGQIN SIGN SLOW FOUR, based on document L2/24-071R and Section 38 of document L2/24-067.

39) L2/24-078: Preliminary Proposal to Add Standardised Variation Sequences for CJK Strokes [IRG N2681]

Comments: The WG reviewed and discussed this document, which is a proposal to add 17 Standardized Variation Sequences (SVSes) for characters in the [CJK Strokes](#) block. The WG

agreed that additional evidence of use in “plain text” would be helpful. More importantly, the WG felt that the proposed SVS names are inconsistent when compared with their Chinese names as found in the provided evidence, and could be improved by simply making them descriptive, such as the following five examples:

Proposed:

```
31C0 FE00; CJK stroke DT form; # CJK STROKE T
31C6 FE00; CJK stroke HWG form; # CJK STROKE HZG
31C6 FE01; CJK stroke XHZG form; # CJK STROKE HZG
31D0 FE00; CJK stroke XH form; # CJK STROKE H
31D2 FE00; CJK stroke BP form; # CJK STROKE P
```

Improved:

```
31C0 FE00; CJK stroke T dotted form; # CJK STROKE T
31C6 FE00; CJK stroke HZG curved form; # CJK STROKE HZG
31C6 FE01; CJK stroke HZG sloping form; # CJK STROKE HZG
31D0 FE00; CJK stroke H sloping form; # CJK STROKE H
31D2 FE00; CJK stroke P flat form; # CJK STROKE P
```

The WG also noted that adding SVSes for these CJK strokes could potentially preclude their encoding as separate characters in the future.

Recommendations: The WG recommends to the UTC the following disposition ⇒

Action Item for Ken Lunde, CJK: Convey to the author of document L2/24-078 the CJK & Unihan Working Group comments in Section 39 of document L2/24-067.

40) L2/24-096: Proposal to add Sichuan Yi (ii) indexing to Unicode

Comments: The WG reviewed and discussed this document, which is a proposal to add to the UCD a new data file that provides metadata for the characters in the [Yi Syllables](#) (1,165) and [Yi Radicals](#) (55) blocks. The proposal author was present during the meeting, and led the discussion. The Script Encoding Working Group will also be reviewing this document.

The WG agreed that the reading data could benefit from a better prefix, such as **kYi_Pinyin** instead of **kSY_Pinyin**, and the reading data itself, which includes the dictionary heading after a period, could be simplified by either uppercasing the dictionary heading portion, or inserting a hyphen after it, such as **BBix** or **bb-ix** instead of **bbix.BB**. Some WG members checked some books after the meeting, and reported that the order and collation appear to be stable and correct. For the other three properties, using the **kSY_** prefix is okay.

Some of the WG members were uncomfortable with the name of the **kSY_VariantOf**, because the relationship is not exactly that of a variant in the usual sense, so the proposal author is encouraged to consider a different name for this particular property.

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

Other Topics

41) New Working Group Name

Comments: The [working groups](#) that report to the UTC were recently renamed to include “Working Group” in their names, and for this working group, this simply meant that “Working” was inserted into the name: **CJK & Unihan Working Group**. Its abbreviation, **CJK**, is unchanged. In addition, the [Unihan Mailing List](#) is considered the official mailing list of this working group.

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

42) Stale UTC Action Items

Comments: The chair of the UTC asked each of the Working Group chairs to report on the status of any [Open Action Items](#) assigned to their WG that are older than five years. For this WG, only [Action Item 159-A56](#) comes close, which will become five years old less than a week after the UTC #179 meeting. The chair of this WG reached out to Richard Cook on 2024-03-31 and again on 2024-04-07 to get the status of this perpetually “in progress” open action item, which will be closed if no response is received by the end of the UTC #179 meeting.

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

43) IRG Working Set 2024 Submission Status

Comments: Including the 89 additions to UAX #45 from Sections 24 through 31 of this document, the number of candidate ideographs in the current draft of the UTC’s submission for IRG Working Set 2024 (aka document IRG N2687), whose submission deadline is **2024-07-26**, is **249**. Among the 249 candidates, there are only five—UTC-01256, UTC-01272, UTC-01276, UTC-03105, and UTC-03118—that may not have sufficient evidence, and are therefore subject to removal prior to submission. The first three of those five UAX #45 ideographs were withdrawn from IRG Working Set 2015 (aka Extension G), and the last two were withdrawn from IRG Working Set 2017 (aka Extension H).

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

44) Updated Provisional Unihan Database Properties

Comments: The purpose of this section is to provide a synopsis of adjustments that were made to provisional Unihan database properties by Ken Lunde since the UTC #178 meeting, which are shown in the table below:

Property	Additions	Removals	Changes	Notes
kPhonetic	197	21	232	Fine-tuning of property values with stakeholder

Recommendations: The WG recommends to the UTC the following disposition ⇒

No UTC action is required.

That is all.