

Title: CJK & Unihan Group Recommendations for UTC #169 Meeting

Author: Ken Lunde

Date: 2021-10-09 (L2/21-173: 2021-09-29)

This document provides the CJK & Unihan Group recommendations for UTC #169 based on a meeting that took place from 6 to 9PM PDT on 2021-09-24, which was attended by Yi Bai, Eiso Chan, Peter Constable, John Jenkins, Ken Lunde, William Nelson, Xieyang Wang, Yifán Wáng, and Haocun Yu via [Zoom](#). The Chair of the [CJK & Unihan Group](#), Ken Lunde, chaired the meeting with John Jenkins as Vice-Chair.

The CJK & Unihan Group reviewed public feedback and documents that were received since UTC #168. Comments are marked in **green**, and Copy&Paste-ready Recommendations to the UTC are marked in **red**.

1) Public Feedback

The seven items of public feedback from [L2/21-169](#) that are in this section were discussed by the group.

Date/Time: Tue Jul 20 00:42:55 CDT 2021

Name: Jerome Alan Rossignuolo

Report Type: Other Question, Problem, or Feedback

Opt Subject: **Missing Fundamental Chinese BuShou Radicals in UNICODE**

Hello,

I am doing research into a learning tool for Simplified Chinese. I find it exceedingly curious that of the 280 Primary and Associated Indexing Components (部首) as standardized by GB13000.1 Chinese Character Component Standard (see <http://ling.whu.edu.cn/law/002/2016-04-20/1307.html>) and used by the XinHuaZiDian ((新华字典) the most widely used dictionary in China, four are missing from UNICODE! That is, there are 276 of these encoded in UNICODE and four have no UNICODE encoding!

(A little note, the term Indexing Components and Radicals are commonly imprecisely interchanged. Most people, even fluent Chinese speakers will incorrectly refer to what are actually Indexing Components as Radicals.)

Moreover, it is obviously not just me faced with the difficulties of not being able to type these Indexing Components. I have scoured the Internet looking for information on them and cannot find any reference to them in the 100s of documents that do reference the 276 others. These four are simply left out of virtually all material. Although you can find them in a scant few places as images. They are in the PDF files of GB13000.1 and are in the printed XinHuaZiDian dictionary along with they are in the official XinHuaZiDian mobile app.

These Associated Indexing Components are still in use. Why are they missing from UNICODE? The Indexing Components are the fundamental building blocks of Chinese. In a way, they are somewhat analogous to our alphabet. It is like

missing the letter little q from UNICODE. Although the difference is that one can still type a Chinese character in UNICODE that includes the missing Indexing Component. Still, it is exceeding odd these cannot be typed since they need to be typed in any article referencing the Indexing Components in the Chinese language used by billions of people. The lack of them on the Internet is testament this is clearly causing people trouble.

Since I cannot type them, I will refer you to reference the missing Indexing Components in the XinHuaZiDian. They are [50], [55], [68], and [145]. These can also be found in GB13000.1.

If you need additional clarification or information, please contact me at [redacted]. I can send you a photo of the missing Indexing Components.

Sincerely,

Jerome Rossignuolo

Comments: This feedback has already been addressed through a private discussion between the feedback submitter and Ken Lunde, and by virtue of the feedback submitter preparing and submitting [L2/21-149](#), which will be covered in Section 3 in this document.

Date/Time: Mon Aug 2 17:35:50 CDT 2021

Name: William He

Report Type: Error Report

Opt Subject: **Error in definition of U+8561 蕒**

Hello,

The definition of U+8561 蕒 is listed as “hemp seeds; plant with abundant”. Something seems to be missing from this definition. Perhaps it means to say “plant with abundant fruit” or something to that effect.

Thanks,

William

Comments: The group discussed this feedback, and agreed with a suggested change made by John Jenkins. John Jenkins already adjusted the *kDefinition* (provisional) property value of U+8561 蕒 as follows:

Current:

U+8561 kDefinition hemp seeds; plant with abundant

Adjusted:

U+8561 kDefinition luxurious, abundant; hemp seeds

Recommendations (Unihan-UTC169-R01): The group recommends to the UTC the following disposition ⇒ **Action Item for Ken Lunde:** Convey to the feedback submitter the CJK & Unihan Group comments in document [L2/21-173](#). [Mon Aug 2 17:35:50 CDT 2021]

Date/Time: Thu Aug 12 09:45:13 CDT 2021

Name: Ken Lunde

Report Type: Error Report

Opt Subject: **kSpoofingVariant errors?**

The latest version of the JIS X 0208 and JIS X 0213 standards explicitly state that U+53F1 𪗇 is a variant of U+20B9F 𪗇, and can be used in implementations that support only the former standard. This is reflected in their kJoyoKanji property values:

U+53F1 kJoyoKanji U+20B9F

U+20B9F kJoyoKanji 2010

The following are the kSpoofingVariant property values for these two ideographs:

U+53F1 kSpoofingVariant U+20B9F

U+20B9F kSpoofingVariant U+53F1

Based on their treatment in Jōyō Kanji and the JIS standards as explicit variants, perhaps they should instead have kZVariant property values:

U+53F1 kZVariant U+20B9F

U+20B9F kZVariant U+53F1

Of course, this is not urgent, and should be considered for Unicode Version 15.0.

Regards...

-- Ken

Comments: The group discussed this feedback, and agreed with the suggested property value changes. John Jenkins already adjusted the provisional *kSpoofingVariant* and *kZVariant* property values of U+53F1 𪗇 and U+20B9F 𪗇 as follows:

Removed:

U+53F1 kSpoofingVariant U+20B9F

U+20B9F kSpoofingVariant U+53F1

Added:

U+53F1 kZVariant U+20B9F

U+20B9F kZVariant U+53F1

Date/Time: Mon Sep 20 23:43:06 CDT 2021

Name: Eiso Chan

Report Type: Error Report

Opt Subject: **4 missing UK glyphs in Unicode, 14.0.0**

The following UK glyphs are missing in Unicode, 14.0.0.

UK-02830 for U+238A7

UK-02849 for U+2A909

UK-01320 for U+2B92E

UK-01422 for U+2E66E

Comments: This issue has already been discussed offline, and will be fixed in the Unicode Version 15.0 code charts.

Date/Time: Thu Sep 23 15:18:26 CDT 2021

Name: Jaemin Chung

Report Type: Website Problem

Opt Subject: **Unihan Database contents search feature update suggestion**

I suggest that the Unihan Database contents search feature be updated.

<http://unicode.org/charts/unihansearch.html>

1. It should not be limited to the definition, Cantonese, Mandarin, Tang, Japanese on/kun, and Korean (Yale).
2. There should be something like “match whole words only” feature. Someone searching for the reading “han” may not want “chang”.
3. For Mandarin, tone numbers no longer work because the Unihan DB now uses tone marks. So the “jing3” example on the main search page should be changed.

Comments: The group discussed this feedback, and agreed that these enhancements to the [Search the Unihan Database](#) page are worth exploring.

Recommendations (Unihan-UTC169-R02): The group recommends to the UTC the following disposition ⇒ **Action Item for John Jenkins, Editorial Committee:** Explore ways to enhance the search feature of the “Search the Unihan Database” page per feedback dated [Thu Sep 23 15:18:26 CDT 2021] in L2/21-169 and Unihan-UTC169-R02 in document L2/21-173, for Unicode Version 15.0. **Action Item for Ken Lunde:** Convey to the feedback submitter the CJK & Unihan Group comments in document L2/21-173. [Thu Sep 23 15:18:26 CDT 2021]

Date/Time: Thu Sep 23 23:54:38 CDT 2021

Name: Jaemin Chung

Report Type: Error Report

Opt Subject: **U+6F55 and U+23C98**

The following `kSimplifiedVariant` and `kTraditionalVariant` values should be added to the Unihan Database.

U+6F55 `kSimplifiedVariant` U+23C98

U+23C98 `kTraditionalVariant` U+6F55

(U+6F55 is 濰 and U+23C98 is 洸)

Comments: The group discussed this feedback, and agreed that the proposed additions to the provisional `kSimplifiedVariant` and `kTraditionalVariant` properties are appropriate. John Jenkins already applied the following additions to the Unihan database:

U+6F55 `kSimplifiedVariant` U+23C98

U+23C98 `kTraditionalVariant` U+6F55

Recommendations (Unihan-UTC169-R03): The group recommends to the UTC the following disposition ⇒ **Action Item for Ken Lunde:** Convey to the feedback submitter the CJK & Unihan Group comments in document L2/21-173. [Thu Sep 23 23:54:38 CDT 2021]

Date/Time: Fri Sep 24 00:26:38 CDT 2021

Name: Jaemin Chung

Report Type: Error Report

Opt Subject: **U+44E8 & U+7F43 and U+6C84 & U+6F90**

Here are additional `kSimplifiedVariant` and `kTraditionalVariant` values that should be added to the Unihan Database.

U+44E8 `kTraditionalVariant` U+7F43

U+7F43 `kSimplifiedVariant` U+44E8

(U+44E8 is 薔 and U+7F43 is 薔)

U+6C84 `kTraditionalVariant` U+6F90

U+6F90 `kSimplifiedVariant` U+6C84

(U+6C84 is 漚 and U+6F90 is 漚)

Comments: The group discussed this feedback, and agreed that the proposed additions to the provisional `kSimplifiedVariant` and `kTraditionalVariant` properties are appropriate. John Jenkins already applied the following additions to the Unihan database:

U+44E8 `kTraditionalVariant` U+7F43

U+6C84 `kTraditionalVariant` U+6F90

U+6F90 `kSimplifiedVariant` U+6C84

U+7F43 `kSimplifiedVariant` U+44E8

Recommendations (Unihan-UTC169-R04): The group recommends to the UTC the following disposition ⇒ **Action Item for Ken Lunde:** Convey to the feedback submitter the CJK & Unihan Group comments in document L2/21-173. [Fri Sep 24 00:26:38 CDT 2021]

2) UAX #38 / Unihan Database Documents

The documents in this section were discussed by the group.

L2/21-118R: Preliminary proposal to add a new provisional `kIDS` property (Unihan) (IRG N2492)

L2/21-161: Concerns on the copyright of the IDS data referred by L2/21-118R

L2/21-165: Feedback on IRGN2492 and the preliminary encoding method of early Chinese organic chemical character, Sanskrit transcription, Tibetan transcription, Tangut transcription and Jianzi Musical Notation

Comments: The group discussed these three related documents. There was general agreement that the five IDCs (*Ideographic Description Characters*) in the preliminary proposal are useful and should be considered for encoding after the formal proposal has been submitted, but that the one named *IDEOGRAPHIC DESCRIPTION CHARACTER STROKE SUBTRACTION* should be renamed *IDEOGRAPHIC DESCRIPTION CHARACTER SUBTRACTION* (the word *STROKE* is removed) and should therefore allow components to be subtracted in addition to strokes. In terms of the to-be-proposed `kIDS` property, the group discussed the need to explicitly state

that IDSeS (*Ideographic Description Sequences*) are subject to change as new ideograph components are encoded that can serve to make IDSeS more efficient (aka shorter). In addition, while Section 18.2 of the Core Specification, *Ideographic Description Characters*, already includes language that states that the shortest possible IDS is preferred, Annex I of ISO/IEC 10646 lacks such language.

In terms of the name and location of the new CJK Unified Ideographs block for encoding ideograph components, there was general agreement in the group that *CJK Unified Ideographs Supplement* is a better name than the one suggested in L2/21-118R, which is *CJK Unified Ideographs Components*, and that the unassigned block between the [CJK Compatibility Ideographs Supplement](#) and the end of Plane 2—U+2FA20 through U+2FFFD (1,502 code points)—is a suitable location for such a block.

With regard to the copyright issue that is mentioned in L2/21-161, Ken Lunde will start a discussion with Andrew West about the copyright status of the [IDS.TXT](#) data file.

The group further recommends that L2/21-165 be broken into two separate documents. The first two sections serve as meaningful feedback for L2/21-118R, but the third section is completely unrelated, hence the recommendation. After the third section has been submitted as a new document, the group recommends that it be forwarded to the Script Ad Hoc Group for review.

Given the amount of work that is necessary for researching ideograph components, the targets for encoding the five new IDCs and ideograph components and adding the *kIDS* property to the Unihan database have necessarily shifted to Unicode Versions 16.0 (2023) and 17.0 (2024), respectively.

Recommendations (Unihan-UTC169-R05): The group recommends to the UTC the following disposition ⇒ **Action Item for Ken Lunde:** Convey to the author of L2/21-161 the CJK & Unihan Group comments in document L2/21-173.

L2/21-150: Specifying Additional Sources for Unihan Database Values

Comments: The group discussed this proposal, and agreed that being able to specify additional sources for Unihan database property values is a very worthwhile addition to UAX #38. The group also favored the use of the “u” prefix when specifying additional sources.

Recommendations (Unihan-UTC169-R06): The group recommends to the UTC the following disposition ⇒ Authorize a proposed update of UAX #38 for Unicode 15.0 to document the additional sources for Unihan database values, based on document L2/21-150 and Unihan-UTC169-R06 in document L2/21-173. **Action Item for John Jenkins, Editorial Committee:** Produce a proposed update of UAX #38 for Unicode 15.0 to document the additional sources for Unihan database values, based on document L2/21-150 and Unihan-UTC169-R06 in document L2/21-173. **Action Item for Rick McGowan:** Post PRI for UAX #38 for Unicode 15.0, to close 2022-01-10.

L2/21-151: Update to the Syntax for the Unihan Database's `kTotalStrokes` Field

Comments: The group discussed this proposal, and agreed that the two-letter identifier for UK-Source ideographs, **Uk**, should be changed to the single-letter identifier **B**, and that the UTC should discuss the merits of the proposal, keeping in mind that such a syntax change has the potential to break existing parsers that use the `kTotalStrokes` property.

Recommendations (Unihan-UTC169-R07): The group recommends to the UTC the following disposition ⇒ Discuss the merits of the proposal, and consider forwarding to stakeholders to solicit feedback.

L2/21-152: Revised Proposal to Disunify U+5F50 (IRG N2509)

Comments: The group discussed this proposal, which was also discussed during IRG #57. There was no conclusion during IRG #57, because China was not yet able to agree to disunify U+5F50 𠄎 into three separate ideographs: 𠄎, 𠄎, and 𠄎. China and Vietnam were given an action by the IRG to review the proposal then report back during IRG #58.

Recommendations (Unihan-UTC169-R08): The group recommends to the UTC the following disposition ⇒ **Action Item for Ken Lunde:** Convey to the author of L2/21-152 the CJK & Unihan Group comments in document L2/21-173.

3) UAX #45 / U-Source Documents

The documents in this section were discussed by the group, and for convenience, the `USource-changes-20210924.txt` and `USource-additions-20210924.txt` (PDF attachments) files include all of the recommended UAX #45 `USourceData.txt` data file changes and additions based on this document. These recommended data file changes and additions are also shown inline, following the Recommendations. In addition, the `UTCHan2021u169-Regular.ttf` TrueType font (PDF attachment) includes the representative glyphs for adding to the UAX #45 `UTCHan.ttf` TrueType font.

L2/21-149: Proposal to add one ideograph to UAX #45 and to change the status of an existing UAX #45 ideograph

Comments: The group reviewed this proposal, and agreed that changing the UAX #45 status value of UTC-00792 𠄎 from W (*Not suitable for encoding as a CJK Unified Ideograph*) to N (*Planned to be submitted for a future extension*) is reasonable, and that adding the new ideograph to UAX #45 with a status value of N is also reasonable. Both ideographs can therefore be considered for inclusion in the UTC's submission for the next IRG working set, or as ideograph components. The group noted that existing ideographs that include UTC-00792 𠄎 as a left-side component currently use the full form, U+8F66 𠄎, in their IDses, meaning that such IDses can be changed in the future if UTC-00792 is encoded.

Recommendations (Unihan-UTC169-R09): The group recommends to the UTC the following disposition ⇒ Change the UAX #45 status value of UTC-00792 from W to N and accept one new U-Source ideograph as **UTC-03248** with a UAX #45 status value of N, based on document L2/21-149 and Unihan-UTC169-R09 in document L2/21-173, for Unicode Version 15.0. **Action Item for John Jenkins:** Change the UAX #45 status value of UTC-00792 from W to N and add one new record to *USourceData.txt* and its representative glyph to *USourceGlyphs.pdf*, based on document L2/21-149 and Unihan-UTC169-R09 in document L2/21-173, for Unicode Version 15.0. **Action Item for Ken Lunde:** Convey to the author of L2/21-149 the CJK & Unihan Group comments in document L2/21-173.

Current:

UTC-00792;W;;159'.0;;;车;WL EF95;;;

Adjusted:

UTC-00792;N;;159'.0;;;车;WL EF95, UTCDoc L2/21-149 1;;4;0

Add:

UTC-03248;N;;129'.0;;;𠮟一;UTCDoc L2/21-149 2;;6;0

L2/21-178: Proposal to add two characters and change the status field of one character in UAX #45

Comments: The group reviewed this proposal, and agreed that changing the UAX #45 status value of UTC-00777 from X (*No action taken*) to N is reasonable, and that adding two new ideographs to UAX #45 with a status value of N is also reasonable. The three ideographs can therefore be considered for inclusion in the UTC's submission for the next IRG working set.

Recommendations (Unihan-UTC169-R10): The group recommends to the UTC the following disposition ⇒ Change the UAX #45 status value of UTC-00777 from X to N and accept two new U-Source ideographs as **UTC-03249** and **UTC-03250** with a UAX #45 status value of N, based on document L2/21-178 and Unihan-UTC169-R10 in document L2/21-173, for Unicode Version 15.0. **Action Item for John Jenkins:** Change the UAX #45 status value of UTC-00777 from X to N and add two new records to *USourceData.txt* and their representative glyphs to *USourceGlyphs.pdf*, based on document L2/21-178 and Unihan-UTC169-R10 in document L2/21-173, for Unicode Version 15.0. **Action Item for Ken Lunde:** Convey to the author of L2/21-178 the CJK & Unihan Group comments in document L2/21-173.

Current:

UTC-00777;X;;85.10;;;𠮟 兹;kHanYu 31707.010;;12;

Adjusted:

UTC-00777;N;;85.10;;;𠮟 兹;kHanYu 31707.010, KP1-52B4, UTCDoc L2/21-178 1;;13;4

Add:

UTC-03249;N;;120.8;;;𠮟 糸 東;KP1-671B, UTCDoc L2/21-178 2;;14;1

UTC-03250;N;;75.5;;;𠮟 木 朮;KP1-4B26, UTCDoc L2/21-178 3;;9;1

L2/21-179: Proposal to add 𐄂 to UAX #45

Comments: The group reviewed this proposal, and agreed that changing the UAX #45 status value of UTC-01005 𐄂 from U to N is reasonable considering the pending disunification of U+5F50 𐄂, and that adding the new ideograph 𐄂 to UAX #45 with a status value of N is also reasonable for the same reason. Both ideographs can therefore be considered for inclusion in the UTC's submission for the next IRG working set, or as ideograph components.

Recommendations (Unihan-UTC169-R11): The group recommends to the UTC the following disposition ⇒ Change the UAX #45 status value of UTC-01005 from U to N and accept one new U-Source ideograph as **UTC-03251** with a UAX #45 status value of N, based on document L2/21-179 and Unihan-UTC169-R11 in document L2/21-173, for Unicode Version 15.0. **Action Item for John Jenkins:** Change the UAX #45 status value of UTC-01005 from U to N and add one new record to *USourceData.txt* and its representative glyph to *USourceGlyphs.pdf*, based on document L2/21-179 and Unihan-UTC169-R11 in document L2/21-173, for Unicode Version 15.0.

Current:

UTC-01005;U;U+5F50;58.0;;;UTCDoc L2/12-333 37;;;3;

Adjusted:

UTC-01005;N;;;58.0;;𐄂𐄂—;UTCDoc L2/12-333 37, UTCDoc L2/21-179 2;;;3;0

Add:

UTC-03251;N;;;58.0;;𐄂𐄂—;UTCDoc L2/21-179 1;;;3;0

4) IRG & IRG-related Documents

The documents in this section were discussed by the group, and for convenience, the *Unihan-additions-20210924.txt* (PDF attachment) file includes all of the recommended Unihan database additions based on this document. These recommended Unihan database additions are also shown inline, following the Recommendations.

L2/21-148: US/Unicode Activity Report for IRG #57 (IRG N2504)

Comments: This is an FYI document.

L2/21-162: Editorial Report (IRG N2506)

Comments: This is an FYI document, because any UTC actions are covered by separate documents that are referenced in this document, all of which have been cross-posted to the UTC document register.

L2/21-158: IRG Meeting #57 Recommendations and Action Items (IRG N2500)

Comments: Although this is largely an FYI document, because almost all UTC actions are covered by separate documents that are referenced in this document, all of which have been

cross-posted to the UTC document register, the group would like to draw the UTC's attention to the IRG's future meeting schedule, the current status of CJK Unified Ideographs Extension H (aka IRG Working Set 2017), and the current status of IRG Working Set 2021.

The IRG #57 meeting took place from 2021-09-13 through 2021-09-17 via Zoom (virtual), and approximately 50 member body delegates and individual experts attended. The Unicode Consortium was the meeting host.

The IRG's future meeting schedule is as follows:

IRG #58: 2022-03-14 through 2022-03-18 via Zoom (virtual)

IRG #59: 2022-10-17 through 2022-10-21 at Netflix (Los Gatos, CA, USA) with Yale University (New Haven, CT, USA) and Zoom (virtual) as backups

IRG #60: 2023-03-20 through 2023-03-24 via Zoom (virtual)

Going forward, after face-to-face meetings can resume, the IRG plans to have one face-to-face meeting and one Zoom (virtual) meeting per year.

In terms of the status of Extension H, T13-3647 (02229) is to be unified with T13-3643 (02227), and GZ-4322205 (05030), which was mistakenly removed, is to be restored. The number of ideographs is therefore expected to be unchanged at 4,192 (U+31350 through U+323AF) when compared to the draft code chart that was posted to the IRG #57 document register as [IRG N2475](#). The latest draft of the Extension H code chart, which incorporates these minor adjustments, is available in the IRG #57 document register as [IRG N2513](#). The group feels that Extension H is now mature enough to accept for Unicode Version 15.0.

In terms of the status of IRG Working Set 2021, the working schedule to produce Version 2.0 (IRG N2516) is as follows:

2021-10-15: The IRG Chief Editor fixes all questionable attributes (excluding radical issues) through ORT (*Online Review Tool*). Updated fonts are provided to the IRG Chief Editor and ORT Manager.

2021-10-22: The ORT Manager releases IRG N2516 (aka IRG Working Set 2021 Version 2.0) through the ORT, and distributes the attributes file (Excel).

2021-12-24: IRG editors submit comments through the ORT.

2022-01-07: The ORT Manager releases Consolidated Comments through the ORT and distributes PDF file.

2022-03-04: IRG editors respond to comments through the ORT

Recommendations (Unihan-UTC169-R12): The group recommends to the UTC the following disposition ⇒ Accept 4,192 CJK Unified Ideographs in a new block CJK Unified Ideographs Extension H with code point range U+31350..U+323AF, based on document IRG N2513 and Unihan-UTC169-R12 in document L2/21-173, for Unicode Version 15.0. **Action Item for Michel Suignard, John Jenkins:** Add the required property values for the CJK Unified Ideographs

Extension H block—IRG source references, kRSUnicode, and kTotalStrokes—to the Unihan database, based on document IRG N2513 and Unihan-UTC169-R12 in document L2/21-173, for Unicode Version 15.0. **Action Item for Ken Whistler:** Update the pipeline to include CJK Extension H.

L2/21-159: Request for Horizontal Extension in the H-column of the ISO/IEC 10646 Standard (IRG N2507)

Comments: This document is a two-ideograph horizontal extension from Hong Kong SAR that also means that these ideographs have been added to their HKSCS standard per the [What is HKSCS? \(Chinese\)](#) page. This horizontal extension was accepted by the IRG during the IRG #57 meeting, so it is reasonable for the UTC to accept it as well. Hong Kong SAR was assigned an action item from the IRG to provide an updated font to Michel Suignard by 2021-09-24.

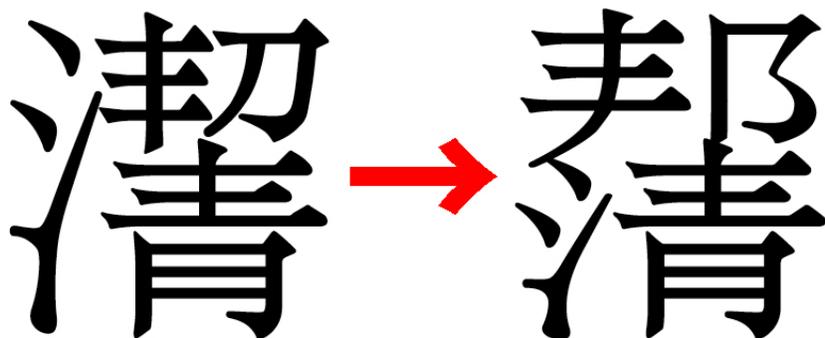
Recommendations (Unihan-UTC169-R13): The group recommends to the UTC the following disposition ⇒ Accept the two-ideograph horizontal extension for U+2BB37 (Extension E) and U+2D25D (Extension F), based on document L2/21-159 (aka IRG N2507) and Unihan-UTC169-R13 in document L2/21-173, for Unicode Version 15.0. **Action Item for Michel Suignard, John Jenkins:** Add the following *kIRG_HSource* property values to the Unihan database, based on document L2/21-159 (aka IRG N2507) and Unihan-UTC169-R13 in document L2/21-173, for Unicode Version 15.0:

Add:

U+2BB37	kIRG_HSource	HD-2BB37
U+2D25D	kIRG_HSource	HD-2D25D

L2/21-160: Glyph issue for U+30759 (IRG N2511)

Comments: The UTC delegates who attended IRG #57, Ken Lunde and John Jenkins, agreed that the representative glyph for U+30759 (UTC-01250) should be adjusted per L2/21-160 as shown below. The UTC was assigned an action item from the IRG to provide an updated font to Michel Suignard by 2021-09-24. As UAX #45 editor, John Jenkins is therefore tasked to do this.



Recommendations (Unihan-UTC169-R14): The group recommends to the UTC the following disposition ⇒ **Action Item for John Jenkins:** Provide to Michel Suignard an updated UTCHan.ttf TrueType font that includes the adjusted form of U+30759 (UTC-01250), based on document L2/21-160 (aka IRG N2511) and Unihan-UTC169-R14 in document L2/21-173, for Unicode Version 15.0.

L2/21-163: Request for Glyph Change in the H-column of the ISO/IEC 10646 Standard (IRG N2508)

Comments: This document from Hong Kong SAR proposes that the H-Source representative glyph for U+22ACF 扌 be changed. The Unicode Version 14.0 code chart representative glyph is shown below. This representative glyph change is also reflected on the [What is HKSCS? \(Chinese\)](#) page. This representative glyph change was accepted by the IRG during the IRG #57 meeting, so it is reasonable for the UTC to accept it as well. Hong Kong SAR was assigned an action item from the IRG to provide an updated font to Michel Suignard by 2021-09-24.

22ACF
手 64.5



H-8ACB

5) Other Documents

The documents in this section were discussed by the group.

L2/21-164: Proposal to revise the description of “kana” in “Han” section in the Unicode Standard-Core specification

Comments: The group reviewed this proposal, which is a from a group that is translating the text of the *Core Specification* into Chinese, and concluded that the first option that is presented, which is to remove altogether the final sentence from the English translation of the 漢字 (*kanji* or *Han characters*) entry in 広辞苑 (*kōjien*) that appears on [page 732](#) of Version 14.0 of the *Core Specification*, results in simpler and less-confusing text. To be clear, the sentence in question is shown below:

They are also called *mana* (true names, as opposed to *kana*, false or borrowed names).

Recommendations (Unihan-UTC169-R15): The group recommends to the UTC the following disposition ⇒ **Action Item for Julie Allen, Editorial Committee:** Remove the final sentence from the English translation of the 漢字 (*kanji* or *Han characters*) entry in 広辞苑 (*kōjien*) that appears on page 732 of Version 14.0 of the *Core Specification*, based on document L2/21-164 and Unihan-UTC169-R15 in document L2/21-173, for Unicode Version 15.0. **Action Item for Ken Lunde:** Convey to the author of L2/21-164 the CJK & Unihan Group comments in document L2/21-173.

L2/21-177: Preliminary Proposal from SAT CJK Buddhist Symbols Collection

Comments: The group reviewed this preliminary proposal, which includes several characters that SAT submitted for IRG Working Set 2021 that the IRG deemed unsuitable for encoding

as CJK Unified Ideographs. The group agreed that some of the characters could potentially be handled as combining sequences that include U+20DD COMBINING ENCLOSING CIRCLE and shaped via the 'ccmp' (*Glyph Composition/Decomposition*) feature, and that others seem worth of encoding, such as in the [Enclosed Ideographic Supplement](#) or [Ideographic Symbols and Punctuation](#) blocks. Due to the largely symbolic nature of these characters, we recommend that the Script Ad Hoc Group review a revised or final version of this proposal.

6) Other Unihan-related Topics

The Unicode Consortium established a formal liaison relationship with the SAT (*SAṃgaṇikīkṛtaṃ Taiśotriṭakam*) Daizōkyō Text Database Committee (aka [SAT](#)). Ken Lunde serves as the Unicode Consortium's liaison contact, and Kiyonori Nagasaki serves as SAT's liaison contact.

In terms of the status of licensing the [Moji Jōhō Kiban Database](#) from CITPC (*Character Information Technology Promotion Council*) to 1) establish the new provisional *kMojiJoho* property as proposed in [L2/20-146](#); 2) improve and expand the existing provisional *kMorohashi* property as proposed in [L2/21-032](#); and 3) establish a new provisional *kJapanese* property that provides Japanese readings expressed in hiragana and katakana, the Unicode Consortium received a draft license agreement from CITPC on 2021-01-23. After consulting with Legal, feedback was provided to CITPC on 2021-03-07. The last communication with CITPC on this matter was on 2021-09-27 (PDT). Due to circumstances beyond our control, these Unihan database additions and enhancements are now targeted for Unicode Version 16.0 (2023).

That is all.