

Title: CJK & Unihan Group Recommendations for UTC #170 Meeting

Author: Ken Lunde

Date: 2022-01-26

This document provides the CJK & Unihan Group recommendations for UTC #170 based on a meeting that took place from 6 to 9PM PST on 2022-01-14, which was attended by Yi Bai, Eiso Chan, Henry Chan, Lee Collins, Peter Constable, John Jenkins, Ken Lunde, William Nelson, Xieyang Wang, and Yifán Wáng 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 #169. Comments are marked in **green**, and Copy&Paste-ready Recommendations to the UTC are marked in **red**.

Public Feedback

The eight items of public feedback from [L2/22-018](#) that are in this section were discussed by the group, and for convenience, the following data files, all of which are PDF attachments, include all of the recommended Unihan database and UAX #45 *USourceData.txt* data file changes based on them

- *Unihan-changes-20220114.txt*
- *kMandarin-changes-20220114.txt*
- *USource-changes-20220114.txt*

With the exception of *kMandarin* property value changes, these recommended changes are also shown inline, following the Recommendations.

01) 2021-10-08 20:34:45 CDT

Date/Time: Fri Oct 8 20:34:45 CDT 2021

Name: Eduardo Marín Silva

Report Type: Other Document Submission

Opt Subject: **Recommended addition of UAX #38**

I would like to apologize if in the past I wasn't as helpful to the Unihan group as I should have.

This is a more succinct recommendation to the Unihan group, to add the number of entries of each field in the description boxes of the document in question: <https://unicode.org/reports/tr38/>. Since Dr. Lunde already compiled the most the number of entries in this document:

https://docs.google.com/spreadsheets/d/1_ad7Z9qqMONlK5SUfNjaSXSNIaG-0P0QKTlUqCvx dhI/edit#gid=559817095

it would be trivial to add the most up to date counts in new versions of UAX #38. This would be convenient for users, that might not want to look at different documents for that info. This info can give a sense to users of how large each field is and comparing the counts in future versions can also reflect the growth of the database.

Comments: The group discussed this feedback, and while there was overwhelming agreement that including this information in UAX #38 has merit, the main concern was about the amount of work that is necessary to keep it up-to-date for each subsequent version of the Unicode Standard and the possibility of introducing errors when doing so. To this end, members of the group are already exploring ways to automate the initial population of this information and subsequent updates through the use of a script, and William Nelson has already demonstrated this with a working Ruby script. The proposed update of UAX #38 for Unicode Version 15.0, [PRI 437](#), is already open with a closing date of 2022-04-10, so the proposed update would simply need to be updated to reflect this new row in each of the Unihan property tables. Below is a mockup that shows the table for the *kTotalStrokes* property:

Property	kTotalStrokes
Status	Informative
Category	IRG Sources
Entries	93,867
Introduced	3.1
Delimiter	space
Syntax	[1-9][0-9]{0,2}
Description	The total number of strokes in the character (including the radical). 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. The preferred value is the one most commonly associated with the character in modern text using customary fonts. This field is targeted specifically for use by CLDR collation and transliteration. As such, it is subject to considerations that help keep pinyin-based Han collation (and its tailorings) and transliteration reasonably stable.

In terms of the [Unihan Property History](#) spreadsheet that Ken Lunde prepared last year, one suggestion from the group is for him to consider drafting a new UTN (*Unicode Technical Note*) that would include the Unihan Property History spreadsheet as a downloadable asset (aka Excel file), which could subsequently be referenced in [Section 5, History](#), of UAX #38.

Recommendations: The group recommends to the UTC the following disposition ➡

Action Item for John Jenkins, Editorial Committee: Update all of the Unihan property tables in Section 4 of UAX #38 by inserting between the rows labeled “Category” and “Introduced” a new row labeled “Entries,” and update PRI 437 accordingly, for Unicode Version 15.0.

Action Item for Ken Lunde: Produce a draft of a new UTN (Unicode Technical Note) that documents the Unihan Property History spreadsheet and includes it as a downloadable asset.

Action Item for Rick McGowan: Convey to the feedback submitter the CJK & Unihan Group comments in Section 01 of document L2/22-022. See feedback [Fri Oct 8 20:34:45 CDT 2021] in document L2/22-018.

02) 2021-11-20 04:41:17 CST

Date/Time: Sat Nov 20 04:41:17 CST 2021

Name: Eiso Chan

Report Type: Error Report

Opt Subject: **Unihan Database**

The kMandarin value for U+266E8 勝 is lao. Maybe we need to modify it to láo based on the corresponding Traditional variant U+6725 勝. This character is very common in Teochow-Swatow Min-dialects for the local food 勝饼, but it's a pity that it has not been included in TGH.

Comments: The group discussed this feedback, and agreed that the modified *kMandarin* property value is reasonable, and should therefore be checked by the CLDR-TC.

Recommendations: The group recommends to the UTC the following disposition ➡

Action Item for Peter Edberg: Ask the CLDR-TC to check the proposed kMandarin property value change for U+266E8, then report back to the UTC. See feedback [Sat Nov 20 04:41:17 CST 2021] in document L2/22-018 and Section 02 of document L2/22-022.

03) 2021-11-27 22:35:35 CST

Date/Time: Sat Nov 27 22:35:35 CST 2021

Name: Jerry Rossignuolo

Report Type: Error Report

Opt Subject: **19227-n5100r-10646-6th-ed-cd3-chart.pdf**

Hello,

I see you have the radical "𠃉" interpreted as the radical "sun". 𠃉 is used as variant of 日 (sun) in the 新华字典部首 (XinHuaZiDian BuShou). Yet not all variants of a BuShou per GF 0011-2009 are of the same radical.

Most of the material I am finding lists the radical 𠃉 is a variant of 冂 with the meaning of cap or hat. I believe this dates back to the Shuowen Jiezi (说文解字) dictionary. <http://www.shuowen.org/?bushou=%E5%86%83>

I also believe 𠃉 as a radical is named 冒字头 which further has me thinking this radical has the meaning of cap or hat.

Yet, I am not sure. Is it possible if you could clarify this for me? I ran across this while working on a Chinese language learning tool and need to correctly identify the meaning of 𠃉.

Thanks,

Jerry Rossignuolo

Comments: This feedback has been resolved offline. Ken Lunde had a private discussion with the feedback submitter, which revealed that this was effectively a request to rename U+2E9C 𠃉 CJK RADICAL SUN, but after it was explained that Unicode character names cannot be changed once encoded, the feedback submitter understood. Therefore, no UTC action is necessary.

Below is the code chart excerpt for U+2E9C 𠄎 that shows its informative note and two cross-references, the first of which is to U+5183 𠄎 that means *hat*.

2E9C 𠄎 CJK RADICAL SUN

- actually a form of the radical for hat, despite its resemblance in shape to the radical for sun

→ 5183 𠄎

→ 65E5 日

Recommendations: The group recommends to the UTC the following disposition ➡

No UTC action is required.

04) 2021-12-18 19:28:56 CST

Date/Time: Sat Dec 18 19:28:56 CST 2021

Name: Richard Hsieh

Report Type: Error Report

Opt Subject: **CJK chart 4E30 mixed up**

4E30 丰 having HB1-A4A5 and T1-4464 that are the Traditional Chinese character. The rest of the characters are the Simplified Chinese characters. They are two different characters and cannot be mixed up. Could not come up with the Traditional character for the name of a person and other things because of softwares that carried both at the same time could not tell apart but to placed the Simplified Chinese in placed of the Traditional Chinese character.

Comments: The group discussed this feedback, and agreed that what is being described is a non-problem and is not an isolated case. Prototypical examples of a simplified and traditional ideograph occupying the same code point and with identical, near-identical, or seemingly-identical glyphs are U+51FA 出, U+53F0 台, and U+9AA8 骨 whose code chart excerpts are shown below, along with that of U+4E30 丰:

4E30 2.3	丰	丰	丰	丰	丰	丰
	G0-3761	HB1-A4A5	T1-4464	J13-2E26	K1-622C	V1-4A36
51FA 𠄎 17.3	出	出	出	出	出	出
	G0-3376	HB1-A558	T1-4578	J0-3D50	K0-7573	V1-4C6D
53F0 𠄎 30.2	台	台	台	台	台	台
	G0-4C28	HB1-A578	T1-463A	J0-4266	K0-773B	V1-4E3D
9AA8 骨 188.0	骨	骨	骨	骨	骨	骨
	G0-3947	HB1-B0A9	T1-586C	J0-397C	K0-4D69	V1-6C51

Recommendations: 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 Section 04 of document L2/22-022. See feedback [Sat Dec 18 19:28:56 CST 2021] in document L2/22-018.

05) 2022-01-04 07:17:52 CST

Date/Time: Tue Jan 4 07:17:52 CST 2022

Name: Andrew Christopher West

Report Type: Other Document Submission

Opt Subject: **CJK Ext. H U+31682 (UK-10989)**

L2-21/053 "Additional repertoire for a future version of Unicode (post Unicode 14.0)" lists the proposed code chart for CJK Unified Ideographs Extension H. The character at U+31682 (UK-10989) is indexed as Radical 40 plus one residual stroke, and placed first under radical 40 (亠). This is clearly wrong because: 1) the character does not include radical 40; and 2) the total stroke count is 9. I suggest changing back to the original proposed index of radical 25 plus 7 residual strokes, and reordering after U+31455.

Comments: The group discussed this feedback, and agreed that this is a critical change for Extension H. In addition to changing the *kRSUnicode* property value of UK-10989 from 40.1 to 25.7, the group agreed that an additional *kRSUnicode* property value, 40.6, should be added. The group further agreed that an additional *kRSUnicode* property value, 25.8, should be added to UK-10140 (aka Extension H U+31726). These changes will affect the ordering of 557 ideographs in Extension H in the range U+31456..U+31682.

31682 𠄎
 亠 40.1 UK-10989

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept the *kRSUnicode* property value changes for UK-10140 and UK-10989, and corresponding changes to code positions of 557 characters in the Extension H block in the range U+31456..U+31682, based on feedback [Tue Jan 4 07:17:52 CST 2022] in L2/22-018 and as amended in Section 05 of document L2/22-022, for Unicode Version 15.0.

Action Item for Michel Suignard: Change the *kRSUnicode* property values of UK-10140 and UK-10989, reorder the Extension H block, and provide to John Jenkins an updated Unihan15.txt data file, for Unicode Version 15.0.

Action Item for John Jenkins: Update the Unihan database to reflect the Extension H changes, for Unicode Version 15.0.

Action Item for Ken Lunde: Convey to the feedback submitter the CJK & Unihan Group comments in Section 05 of document L2/22-022. See feedback [Tue Jan 4 07:17:52 CST 2022] in document L2/22-018.

Current:

U+31682	kRSUnicode	40.1
U+31726	kRSUnicode	50.7

Adjusted:

U+31456	kRSUnicode	25.7	40.6
U+31726	kRSUnicode	50.7	25.8

06) 2022-01-06 07:53:34 CST

Date/Time: Thu Jan 6 07:53:34 CST 2022

Name: Ivan Panchenko

Report Type: Error Report

Opt Subject: **Unihan_Readings.txt**

I want to report some mistakes in the Unihan Database definitions (kDefinition): “from from” (instead of “from”), “disturbe” (instead of “disturb”), “pon your mind” (instead of “on your mind”), “thron” (instead of “thorn”), “flaten”(instead of “flatten”), “name name” (instead of “name”), “chrysanthemun” (instead of “chrysanthemum”), “purpurca” (instead of “purpurea”), “force fo arms” (instead of “force of arms”), “phtholein” (instead of “phthalein”), “the the” (instead of “the”), “ber eaten” (instead of “be eaten”), “bured” (instead of “buried”), “askewd” (instead of “askew”), “foll of cloth” (instead of “roll of cloth”), “smilling” (instead of “smiling”), “without friends or relativ” (instead of “without friends or relatives”), “longtum” (instead of “longum”), “themedia forskali” (instead of “Themeda forskalii”), “circium” (instead of “Cirsium”), “bracenia” (instead of “Brasenia”), “artemesia” (instead of “artemisia”), “corp of a bird” (instead of “crop of a bird”), “stellariana” (instead of “stelleriana”), “eumenes polifomis” (apparently should be “Eumenes pomiformis”, but is this definition actually correct??), “loquatious” (instead of “loquacious”), “interprete” (instead of “interpret”), “liesure” (instead of “leisure”), “mischevious” (instead of “mischievous”), “fy” (instead of “fry”), “incorruptable” (instead of “incorruptible”), “repse” (instead of “repose”).

The following words should be capitalized: “sanskrit”, “buddhist”, “daoist”, “pekinese”, “persian”. Scientific names are capitalized as well, so this should be corrected: “malus”, “canis”, “ursus”, “rubia”, “plantago”, “piper”, “caryopteris”, “hydropyrum”, “artemisia stelleriana”, “gracilaria”, “vitis”, “valeriana”, “pteris”, “ligusticum”, “allium”, “cyperus”, “lophanthus”, “arca”, “libellulidae”, “vipera”, “brachyura”, “acrida”, “cosmopsaltria”, “parasilurus”, “spheroides”, “coryphaena”, “pagrosomus”, “treron”, “grus”.

This is questionable, but I am unsure what this is supposed to mean: “leucacene”, “suffle”.

Comments: The group discussed this feedback, and agreed that these all are worthwhile *kDefinition* property value changes. Ken Lunde did the work to associate the suggested changes to the affected ideographs, which involved contacting the feedback submitter for clarification. Some of the property value changes had already been made by John Jenkins, so they are not reflected in this document. The group also recommends changing “Jap.” to “Japanese” that appears in the *kDefinition* property values for U+4F3D 伽 and U+4F5B 佛.

Recommendations: The group recommends to the UTC the following disposition ➡

Action Item for John Jenkins: Update the Unihan database, based on feedback [Thu Jan 6 07:53:34 CST 2022] in L2/22-018 and as amended in Section 06 of L2/22-022, for Unicode Version 15.0.

Action Item for Ken Lunde: Convey to the feedback submitter the CJK & Unihan Group comments in Section 06 of document L2/22-022. See feedback [Thu Jan 6 07:53:34 CST 2022] in document L2/22-018.

Current:

U+355B	kDefinition	(ancient form of 友) a friend; a companion; a as-sociate, friendly, to make friends of, brotherly regard
U+3630	kDefinition	(interchangeable 斟) name of a old country; used in name of a place
U+3689	kDefinition	(corrupted form) to make a profit on sell and buy; the profit in business, (interchangeable 贏) a overplus; gain; profit; abundance
U+371C	kDefinition	(ancient form form of 傲) proud; haughty; over-bearing rude
U+38DB	kDefinition	walking, (ancient form form of 會) to meet, to gather, to be able, to realize
U+39D4	kDefinition	to disturb; to agitate; to harass, to quarrel; to wrangle
U+3A9C	kDefinition	(same as 揣) to measure; to estimate; to suppose; to gauge what's pon another's mind; to presume; to test; to sound out, (same as 捶) to weigh in the hand
U+3C38	kDefinition	(a abbreviated form of 斂 斂) to collect; to gather; to hold together, to desire
U+4070	kDefinition	to look in a evil or wicked way; wicked vision
U+4079	kDefinition	(same as U+7764 睥) to look askance -- a expres-sion of disdain or despise
U+41B7	kDefinition	a hole; an opening; a aperture; deep; far and profound
U+42B9	kDefinition	(a abbreviated form of 纖) small, fine, delicate
U+434E	kDefinition	a spindle, a earthenware used to pinch or knead, a small crock
U+43AD	kDefinition	a kind of farm tool made of twisted thron strips used to flaten the land
U+44B2	kDefinition	another name name for 黃岑 (Scutellaria baikalen-sis)
U+456E	kDefinition	(interchangeable 菊) flower of the chrysanthemum; chrysanthemum
U+4582	kDefinition	an edible water-plant -- Brasenia purpurca
U+45CF	kDefinition	name of a insect; a black insect with six legs, capable of ejecting poison for self-defense; the spider millipede
U+46AF	kDefinition	(ancient form of 討) to quell (uprising, rebel-lion, etc.) to punish (another nation, etc.) by force fo arms, sound; voice; tone
U+475E	kDefinition	name of a immortal; with man's face and bird's body; be known as (禺號)
U+479B	kDefinition	dancing of a eastern minority group in old China

U+48ED	kDefinition	an organic compound; phtholein
U+49B2	kDefinition	eaves of a ancestral temple (of the ruling family), door of a shrine or a temple, small opening (of a door), to spy; to peep
U+49FC	kDefinition	(same as 鵞) the the snow goose; swan, standing erect; standing quietly, target; bull's-eye
U+4B26	kDefinition	food items merely arranged for display purpose only; not to ber eaten; to set out fruit, etc., in plates arranged only for show
U+4CA2	kDefinition	(simplified form of 鰻) a kind of fish; stout; bluish gray color; big flat head; big mouth and small eyes; usually stay in the bottom of the shallow sea; half bured in the muddy sand
U+4CD9	kDefinition	a egret-like bird, a kind of water bird
U+4F3D	kDefinition	transcription of sanskrit 'gha' in buddhist texts ('samgha', etc.); (nursing; attending; entertainer) (Jap.); temple; in Chinese this character is not used alone
U+4F5B	kDefinition	Buddha; of Buddhism; merciful person; Buddhist image; the dead (Jap.)
U+52C5	kDefinition	imperial decree; daoist magic
U+52D0	kDefinition	imperial degree; daoist magic
U+6877	kDefinition	rafter; malus toringo
U+6B6A	kDefinition	slant; inclined; askewd, awry
U+72C6	kDefinition	pekinese dog, lap dog, pug; pup
U+7444	kDefinition	a ornamental piece of jade several inches across with a hole in the center
U+7550	kDefinition	to fill; a foll of cloth
U+80B8	kDefinition	spread out; smiling or laughing
U+8186	kDefinition	the corp of a bird or fowl; fat
U+8315	kDefinition	alone; without friends or relativ
U+8359	kDefinition	plantago major
U+835C	kDefinition	species of bean; piper longtum
U+83C5	kDefinition	coarse grass, themedia forskali
U+84DF	kDefinition	circium, thistles; surname
U+84F7	kDefinition	motherwort, various plants of the genus Leonurus, such as L. sibiricus and L. artemesia
U+84FD	kDefinition	species of bean; piper longtum
U+8551	kDefinition	valeriana villosa, climbing plant
U+8557	kDefinition	leucacene
U+858A	kDefinition	circium, thistles; surname
U+8629	kDefinition	Artemisia stellariana
U+863A	kDefinition	gracilaria verrucosa
U+8761	kDefinition	eumenes polifomis, kind of wasp
U+8770	kDefinition	vipera russelii siamensis
U+8815	kDefinition	eumenes polifomis, kind of wasp
U+8B27	kDefinition	loquacious; banter, joking
U+8C5A	kDefinition	small pig, suckling pig; suffle
U+91C8	kDefinition	interprete, elucidate; release
U+91CA	kDefinition	interprete, elucidate; release
U+91CB	kDefinition	interprete, elucidate; release
U+9592	kDefinition	liesure; idle; peaceful, tranquil, calm
U+9995	kDefinition	bread (persian naan)
U+9B48	kDefinition	mischevious, one footed spirit that dwells in the mountains

U+9B8E	kDefinition	sheatfish, parasilurus asotus
U+9BAD	kDefinition	salmon; spheroides vermicularis
U+9BD5	kDefinition	coryphaena hippurus
U+9BDB	kDefinition	pagrosomus major, porgy
U+9BE4	kDefinition	spawn; roe; fy
U+9C87	kDefinition	sheatfish, parasilurus asotus
U+9C91	kDefinition	salmon; spheroides vermicularis
U+9CAF	kDefinition	coryphaena hippurus
U+9CB2	kDefinition	spawn; roe; fy
U+9CB7	kDefinition	pagrosomus major, porgy
U+9E58	kDefinition	a kind of pigeon; treron permagna
U+9E73	kDefinition	crane, grus japonensis; stork; Ciconia species (various)
U+20533	kDefinition	economical; incorruptable
U+21A4B	kDefinition	repse, serenity, peace; peaceful; Nanking
Adjusted:		
U+355B	kDefinition	(ancient form of 友) a friend; a companion; an associate, friendly, to make friends of, brotherly regard
U+3630	kDefinition	(interchangeable 斟) name of an old country; used in name of a place
U+3689	kDefinition	(corrupted form) to make a profit on sell and buy; the profit in business, (interchangeable 贏) an overplus; gain; profit; abundance
U+371C	kDefinition	(ancient form of 傲) proud; haughty; overbearing rude
U+38DB	kDefinition	walking, (ancient form of 會) to meet, to gather, to be able, to realize
U+39D4	kDefinition	to disturb; to agitate; to harass, to quarrel; to wrangle
U+3A9C	kDefinition	(same as 揣) to measure; to estimate; to suppose; to gauge what's on another's mind; to presume; to test; to sound out, (same as 捶) to weigh in the hand
U+3C38	kDefinition	(an abbreviated form of 歛 斂) to collect; to gather; to hold together, to desire
U+4070	kDefinition	to look in an evil or wicked way; wicked vision
U+4079	kDefinition	(same as U+7764 睥) to look askance -- an expression of disdain or despise
U+41B7	kDefinition	a hole; an opening; an aperture; deep; far and profound
U+42B9	kDefinition	(an abbreviated form of 纖) small, fine, delicate
U+434E	kDefinition	a spindle, an earthenware used to pinch or knead, a small crock
U+43AD	kDefinition	a kind of farm tool made of twisted thorn strips used to flatten the land
U+44B2	kDefinition	another name for 黃岑 (Scutellaria baikalensis)
U+456E	kDefinition	(interchangeable 菊) flower of the chrysanthemum; chrysanthemum
U+4582	kDefinition	an edible water-plant -- Brasenia purpurea
U+45CF	kDefinition	name of an insect; a black insect with six legs, capable of ejecting poison for self-defense; the spider millipede

U+46AF	kDefinition	(ancient form of 討) to quell (uprising, rebellion, etc.) to punish (another nation, etc.) by force of arms, sound; voice; tone
U+475E	kDefinition	name of an immortal; with man's face and bird's body; be known as (禺號)
U+479B	kDefinition	dancing of an eastern minority group in old China
U+48ED	kDefinition	an organic compound; phthalein
U+49B2	kDefinition	eaves of an ancestral temple (of the ruling family), door of a shrine or a temple, small opening (of a door), to spy; to peep
U+49FC	kDefinition	(same as 鵠) the snow goose; swan, standing erect; standing quietly, target; bull's-eye
U+4B26	kDefinition	food items merely arranged for display purpose only; not to be eaten; to set out fruit, etc., in plates arranged only for show
U+4CA2	kDefinition	(simplified form of 鰐) a kind of fish; stout; bluish gray color; big flat head; big mouth and small eyes; usually stay in the bottom of the shallow sea; half buried in the muddy sand
U+4CD9	kDefinition	an egret-like bird, a kind of water bird
U+4F3D	kDefinition	transcription of Sanskrit 'gha' in Buddhist texts ('samgha', etc.); nursing; attending; entertainer (Japanese); temple; in Chinese this character is not used alone
U+4F5B	kDefinition	Buddha; of Buddhism; merciful person; Buddhist image; the dead (Japanese)
U+52C5	kDefinition	imperial decree; Daoist magic
U+52D0	kDefinition	imperial degree; Daoist magic
U+6877	kDefinition	rafter; Malus toringo
U+6B6A	kDefinition	slant; inclined; askew, awry
U+72C6	kDefinition	Pekinese dog, lap dog, pug; pup
U+7444	kDefinition	an ornamental piece of jade several inches across with a hole in the center
U+7550	kDefinition	to fill; a roll of cloth
U+80B8	kDefinition	spread out; smiling or laughing
U+8186	kDefinition	the crop of a bird or fowl; fat
U+8315	kDefinition	alone; without friends or relatives
U+8359	kDefinition	Plantago major
U+835C	kDefinition	species of bean; Piper longum
U+83C5	kDefinition	coarse grass, Themedia forskalii
U+84DF	kDefinition	Cirsium, thistles; surname
U+84F7	kDefinition	motherwort, various plants of the genus Leonurus, such as L. sibiricus and L. artemisia
U+84FD	kDefinition	species of bean; Piper longum
U+8551	kDefinition	Valeriana villosa, climbing plant
U+8557	kDefinition	genus Leucaena; butterbur; Petasites japonicus
U+858A	kDefinition	Cirsium, thistles; surname
U+8629	kDefinition	Artemisia stelleriana
U+863A	kDefinition	Gracilaria verrucosa
U+8761	kDefinition	genus Eumenes, potter's wasps
U+8770	kDefinition	Daboia siamensis, eastern Russell's viper
U+8815	kDefinition	to squirm, wriggle
U+8B27	kDefinition	loquacious; banter, joking
U+8C5A	kDefinition	small pig, suckling pig; shuffle one's feet
U+91C8	kDefinition	interpret, elucidate; release

U+91CA	kDefinition	interpret, elucidate; release
U+91CB	kDefinition	interpret, elucidate; release
U+9592	kDefinition	leisure; idle; peaceful, tranquil, calm
U+9995	kDefinition	bread (Persian naan)
U+9B48	kDefinition	mischievous, one footed spirit that dwells in the mountains
U+9B8E	kDefinition	sheatfish, <i>Parasilurus asotus</i>
U+9BAD	kDefinition	salmon; <i>Spheroides vermicularis</i>
U+9BD5	kDefinition	<i>Coryphaena hippurus</i>
U+9BDB	kDefinition	<i>Pagrosomus major</i> , porgy
U+9BE4	kDefinition	spawn; roe; fry
U+9C87	kDefinition	sheatfish, <i>Parasilurus asotus</i>
U+9C91	kDefinition	salmon; <i>Spheroides vermicularis</i>
U+9CAF	kDefinition	<i>Coryphaena hippurus</i>
U+9CB2	kDefinition	spawn; roe; fry
U+9CB7	kDefinition	<i>Pagrosomus major</i> , porgy
U+9E58	kDefinition	a kind of pigeon; <i>Treron permagna</i>
U+9E73	kDefinition	crane, <i>Grus japonensis</i> ; stork; <i>Ciconia</i> species (various)
U+20533	kDefinition	economical; incorruptible
U+21A4B	kDefinition	repose, serenity, peace; peaceful; Nanking

07) 2022-01-06 20:28:54 CST

Date/Time: Thu Jan 6 20:28:54 CST 2022

Name: Ken Lunde

Report Type: Error Report

Opt Subject: **Unihan Database**

The `kTotalStrokes` property value of U+2AB8F 𢦏 (𠬞手思) should be changed from 12 to 13, because its indexing radical is composed of four strokes, not three.

Comments: The group discussed this feedback, and agreed with the suggested `kTotalStrokes` property value change for U+2AB8F whose current code chart excerpt is shown below:

2AB8F 𢦏
手 64.9
V4-4937

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept the `kTotalStrokes` property value change for U+2AB8F, based on feedback [Thu Jan 6 20:28:54 CST 2022] in L2/22-018 and Section 07 of document L2/22-022, for Unicode Version 15.0.

Action Item for Michel Suignard: Change the `kTotalStrokes` property value for U+2AB8F from 12 to 13, and provide to John Jenkins an updated Unihan15.txt data file, for Unicode Version 15.0.

Action Item for John Jenkins: Update the Unihan database to reflect the `kTotalStrokes` property value change for U+2AB8F, for Unicode Version 15.0.

Current:

U+2AB8F kTotalStrokes 12

Adjusted:

U+2AB8F kTotalStrokes 13

08) 2022-01-10 18:45:29 CST

Date/Time: Mon Jan 10 18:45:29 CST 2022

Name: Jaemin Chung

Report Type: Error Report

Opt Subject: **Defect report on USourceData.txt**

In USourceData.txt, some IDSes have semicolons in them. This is bad because a semicolon is already used as a delimiter.

Here are the IDSes with semicolons:

UTC-03134;B;U+28559;162.9;;𠄎𠄎&P7-03;刀;UTCDoc L2/17-204;;13 12;3

UTC-03143;WS-2017;;145.12;;𠄎𠄎&H5-01;一石衣;UTCDoc L2/17-204;;18;1

UTC-03156;WS-2017;;94.8;;𠄎&H8-01;犬;UTCDoc L2/17-204;;12;1

Comments: The group discussed this feedback, and agreed that U+FF1B FULLWIDTH SEMICOLON is an appropriate replacement for semicolons when they appear in IDSes (Field 5 of the UAX #45 data file, *USourceData.txt*) as character entity references (similar to HTML character entities) that begin with an ampersand and terminate with a semicolon. These character entity references correspond to an unencoded ideograph component, and are meant for IRG (*Ideographic Research Group*) use.

In addition, the group noted that per the comments in [IRG Working Set 2017 #03966](#) (aka UTC-03143), the IDS for UTC-03143 can be changed to 𠄎𠄎一𠄎石衣, which does not include a semicolon. Furthermore, Fields 1 and 2 of UTC-03143 and UTC-03156 should be updated to reflect their Extension H status and Extension H code points.

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept the UAX #45 data file changes, based on feedback [Mon Jan 10 18:45:29 CST 2022] in L2/22-018 and as amended in Section 08 of document L2/22-022, for Unicode Version 15.0.

Action Item for John Jenkins, Editorial Committee: Document the use of U+FF1B FULLWIDTH SEMICOLON in Field 5 of the UAX #45 data file, based on feedback [Mon Jan 10 18:45:29 CST 2022] in L2/22-018 and as amended in Section 08 of document L2/22-022, for Unicode Version 15.0.

Action Item for John Jenkins: Update the UAX #45 data file, USourceData.txt, based on feedback [Mon Jan 10 18:45:29 CST 2022] in L2/22-018 and as amended in Section 08 of document L2/22-022, for Unicode Version 15.0.

Action Item for Ken Lunde: Convey to the feedback submitter the CJK & Unihan Group comments in Section 08 of document L2/22-022. See feedback [Mon Jan 10 18:45:29 CST 2022] in document L2/22-018.

Current:

UTC-03134;B;U+28559;162.9;;𠂇𠂇&P7-03;刀;UTCDoc L2/17-204;;13 12;3
UTC-03143;WS-2017;;145.12;;𠂇𠂇&H5-01;𠂇石衣;UTCDoc L2/17-204;;18;1
UTC-03156;WS-2017;;94.8;;𠂇&H8-01;犬;UTCDoc L2/17-204;;12;1

Adjusted:

UTC-03134;B;U+28559;162.9;;𠂇𠂇&P7-03;刀;UTCDoc L2/17-204;;13 12;3
UTC-03143;H;U+32043;145.12;;𠂇𠂇𠂇石衣;UTCDoc L2/17-204;;18;1
UTC-03156;H;U+31B7F;94.8;;𠂇&H8-01;犬;UTCDoc L2/17-204;;12;1

UAX #38 / Unihan Database Documents

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

- *Unihan-changes-20220114.txt*
- *kMandarin-changes-20220114.txt*
- *Unihan-additions-20220114.txt*
- *kMandarin-additions-20220114.txt*
- *Unihan-removals-20220114.txt*

With the exception of *kMandarin* property value changes and additions, these recommended Unihan database changes, additions, and removals are also shown inline, following the Recommendations.

09) L2/21-226: Proposed Supplement to the Unihan Database's *kTotalStrokes* Field

Comments: The group discussed this proposal to add a new provisional Unihan database property, *kIRGStrokeCount*, which is intended to supplement the existing *kTotalStrokes* property. The group agreed that *kAlternateTotalStrokes* would be a better property name, and that the *Unihan_DictionaryLikeData.txt* data file would be the best home for it. The proposed update of UAX #38 for Unicode Version 15.0, [PRI 437](#), is already open with a closing date of 2022-04-10, so the proposed update would simply need to be updated to reflect the new property, and to add a new section that documents the single-letter IRG source identifiers that are also used in the existing *kIICore* and *kUnihanCore2020* properties.

Ken Lunde also came up with an initial set of data based on J-Source ideographs that have a different number of strokes than what is currently reflected in the *kTotalStrokes* property. The *kRSAdobe_Japan1_6* property, which includes accurate Japanese-specific radical and stroke information, was used to derive this initial set of data. This exercise effectively pinpointed cases of existing *kTotalStrokes* property values that are either incomplete—for example, the *kTotalStrokes* property value of U+9AA8 骨 is 9, but it should be 9 *and* 10 because the G- and T-Source representative glyphs have 9 and 10 total strokes, respectively—or inaccurate—for example, the *kTotalStrokes* property value of U+2A6B2 臚 is 20, but it should be 15. In other words, adding a *kAlternateTotalStrokes* property value to an ideograph should first require that

its existing *kTotalStrokes* property value be reviewed—and changed, if necessary—so that these properties can be in sync. See the code chart excerpts below:

9AA8 骨 188.0	骨	骨	骨	骨	骨	骨
	G0-3947	HB1-B0A9	T1-586C	J0-397C	K0-4D69	V1-6C51
2A6B2 龜 213.4	臄 J4-7E76					

Note: Document L2/21-226 was prepared per UTC Action Item 168-A61.

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept a new provisional Unihan database property, *kAlternateTotalStrokes*, based on L2/21-226 and as amended in Section 09 of document L2/22-022, with the understanding that adding this property value to any ideograph requires that the existing *kTotalStrokes* property value for that ideograph be reviewed—and modified, if necessary—so that these properties can be in sync, and change the *kTotalStrokes* property values of U+9AA8 and U+2A6B2 to “9 10” and 15, respectively, for Unicode Version 15.0.

Action Item for Michel Suignard: Change the *kTotalStrokes* property values of U+9AA8 and U+2A6B2 to “9 10” and 15, respectively, and provide to John Jenkins an updated Unihan15.txt data file, for Unicode Version 15.0.

Action Item for John Jenkins, Editorial Committee: Document the provisional *kAlternateTotalStrokes* property in UAX #38, add a new section after Section 3.8 that documents the single-letter IRG source identifiers, and update PRI 437 accordingly, for Unicode Version 15.0.

Action Item for Ken Lunde: Provide to John Jenkins the working data for the initial set of *kAlternateTotalStrokes* property values that were derived from the *kRSAdobe_Japan1_6* property for the purpose of populating the *kAlternateTotalStrokes* property with an initial, albeit minimal, set of property values, for Unicode Version 15.0.

Action Item for John Jenkins: Update the Unihan database to reflect the *kTotalStrokes* property value changes for U+9AA8 and U+2A6B2, for Unicode Version 15.0.

Action Item for John Jenkins, CJK & Unihan Group: Populate the *kAlternateTotalStrokes* property with an initial, albeit minimal, set of property values, for Unicode Version 15.0.

Current:

U+9AA8	<i>kTotalStrokes</i>	9
U+2A6B2	<i>kTotalStrokes</i>	20

Adjusted:

U+9AA8	<i>kTotalStrokes</i>	9 10
U+2A6B2	<i>kTotalStrokes</i>	15

10) L2/21-228: Request to move the source reference for UK-02830 (IRG N2520)

Comments: The group reviewed this proposal, and agreed that the *kIRG_UKSource* source reference UK-02830 should be moved from U+238A7 𪛗 to U+4DBE 𪛗. The group further recommends that the *kZVariant* property be added to U+4DBE 𪛗 and U+238A7 𪛗.

4DBE
欠 76.6

𪛗
T6-4A3F

238A7
欠 76.6

𪛗
GKX-0567.07

𪛗
T6-3857

𪛗
UK-02830

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept the proposal to move the *kIRG_UKSource* property value UK-02830 from U+238A7 to U+4DBE, based on document L2/21-228 and Section 10 of document L2/22-022, for Unicode Version 15.0.

Action Item for Michel Suignard: Move the *kIRG_UKSource* property value UK-02830 from U+238A7 to U+4DBE, and provide to John Jenkins an updated Unihan15.txt data file, for Unicode Version 15.0.

Action Item for John Jenkins: Update the Unihan database to reflect the property value changes for U+4DBE and U+238A7, based on document L2/21-228 and as amended in Section 10 of document L2/22-022, for Unicode Version 15.0.

Action Item for Ken Lunde: Convey to the author of L2/21-228 the CJK & Unihan Group comments in Section 10 of document L2/22-022.

Removed:

U+238A7	<i>kIRG_UKSource</i>	UK-02830
---------	----------------------	----------

New:

U+4DBE	<i>kIRG_UKSource</i>	UK-02830
U+4DBE	<i>kZVariant</i>	U+238A7
U+238A7	<i>kZVariant</i>	U+4DBE

11) L2/21-236: Horizontal extension or disunification request for 14 Khangxi-characters

Comments: The group reviewed this proposal, and determined that all 14 ideographs are unifiable with existing ideographs. The group also agreed that there is no benefit in adding these ideographs to UAX #45 for the purpose of horizontally extending the ideographs with which they are unifiable, and that a better approach would be to add *Kangxi Dictionary* references to the existing *kKangXi* property values of the ideographs with which they are unifiable.

Recommendations: The group recommends to the UTC the following disposition ➡

Action Item for John Jenkins, CJK & Unihan Group: Use the *kIRG_GSource* and *kIRGKangXi* properties to derive additional *kKangXi* property values.

Action Item for Ken Lunde: Convey to the author of L2/21-236 the CJK & Unihan Group comments in Section 11 of document L2/22-022.

12) L2/22-008: Proposal to change various values in the Unihan Database

Comments: The group reviewed this proposal, and agreed that the new and modified *kMandarin* property values are reasonable, and should therefore be checked by the CLDR-TC. The group agreed to change the proposed new *kSimplifiedVariant* property value for U+6220 to be in code point order, U+53EA U+6220 (was U+6220 U+53EA), and to change the proposed *kTraditionalVariant* property value change for U+53EA to also be in code point order, U+53EA U+6220 U+96BB (was U+53EA U+96BB U+6220). The remaining property value changes and additions were deemed reasonable by the group.

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept the *kRSUnicode* and *kTotalStrokes* property value changes for U+29867, based on L2/22-008 and Section 12 of document L2/22-022, for Unicode Version 15.0.

Action Item for Peter Edberg: Ask the CLDR-TC to check the proposed *kMandarin* property value changes and additions, then report back to the UTC. See document L2/22-008 and Section 12 of document L2/22-022.

Action Item for Michel Suignard: Change the *kRSUnicode* and *kTotalStrokes* property values for U+29867, based on L2/22-008 and Section 12 of document L2/22-022, and provide to John Jenkins an updated Unihan15.txt data file, for Unicode Version 15.0.

Action Item for John Jenkins: Update the Unihan database, based on L2/22-008 and as amended in Section 12 of document L2/22-022, for Unicode Version 15.0.

Action Item for Ken Lunde: Convey to the author of L2/22-008 the CJK & Unihan Group comments in Section 12 of document L2/22-022.

Current:

U+4661	<i>kDefinition</i>	(simplified form) a knot; a button, (dialect) to tie up; fastening; to wear; to slip on, to make a knot by a rope; a piece of string or a ribbon, (same as 繪) to draw (picture)
U+4CA1	<i>kDefinition</i>	(simplified form U+9C0C 鳊) loach
U+53EA	<i>kTraditionalVariant</i>	U+53EA U+96BB
U+8BA0	<i>kCantonese</i>	gaan2
U+9485	<i>kCantonese</i>	gaan2
U+9963	<i>kCantonese</i>	gaan2
U+9FCD	<i>kDefinition</i>	hill; long and narrow highland; used in place names
U+9FCE	<i>kDefinition</i>	(tǎ) place names in Guangdong and Zhejiang; (dá) cobblestone, water gate
U+29867	<i>kRSUnicode</i>	187.0
U+29867	<i>kTotalStrokes</i>	10
U+2C494	<i>kDefinition</i>	for something hard or bulging to press against the body and cause it discomfort

Adjusted:

U+4661	<i>kDefinition</i>	a knot; a button, (dialect) to tie up; fastening; to wear; to slip on, to make a knot by a rope; a piece of string or a ribbon, (same as 繪) to draw (picture)
U+4CA1	<i>kDefinition</i>	(simplified form of U+9C0C 鳊) loach

U+53EA	kTraditionalVariant	U+53EA U+6220 U+96BB
U+8BA0	kCantonese	jin4
U+9485	kCantonese	gam1
U+9963	kCantonese	sik6
U+9FCD	kDefinition	hill; long and narrow highland; used in place names; (Cant.) throw up and scatter (as dust)
U+9FCE	kDefinition	(tǎ) place name in Zhejiang; (dá) cobblestone, water gate, place name in Guangdong
U+29867	kRSUnicode	187.1
U+29867	kTotalStrokes	11
U+2C494	kDefinition	place name in Guangdong; (Cant.) for something hard or bulging to press against the body and cause it discomfort

Removed:

U+364C	kDefinition	(a simplified form)
U+392B	kDefinition	(non-classical form)
U+41B4	kDefinition	(abbreviated form)
U+497B	kDefinition	(simplified form)
U+497C	kDefinition	(simplified form)
U+4984	kDefinition	(simplified form)

New:

U+4DB6	kVietnamese	đè
U+6220	kSimplifiedVariant	U+53EA U+6220
U+7CF9	kCantonese	mik6
U+7E9F	kCantonese	mik6
U+81E4	kSimplifiedVariant	U+81E4 U+30021
U+8A01	kCantonese	jin4
U+9B97	kSimplifiedVariant	U+9FF4
U+9FB0	kCantonese	zi2
U+9FB5	kCantonese	sau2
U+9FBB	kCantonese	lyun4
U+9FBB	kSimplifiedVariant	U+3005C
U+9FC4	kCantonese	loeng4
U+9FCC	kCantonese	loeng4
U+9FCD	kCantonese	pung1
U+9FCD	kCheungBauerIndex	395.05
U+9FDF	kTraditionalVariant	U+9FE0
U+9FE0	kSimplifiedVariant	U+9FDF
U+9FF0	kCantonese	hon4
U+9FF0	kTraditionalVariant	U+2C267
U+9FF1	kCantonese	baa1
U+9FF2	kCantonese	lou4
U+9FF2	kTraditionalVariant	U+27C0E
U+9FF3	kCantonese	to4
U+9FF3	kSimplifiedVariant	U+9FF8
U+9FF4	kCantonese	dung1
U+9FF4	kTraditionalVariant	U+9B97
U+9FF5	kCantonese	zi3
U+9FF5	kTraditionalVariant	U+29DD3
U+9FF6	kCantonese	loeng4
U+9FF6	kTraditionalVariant	U+29DD5
U+9FF7	kCantonese	jam1
U+9FF7	kTraditionalVariant	U+29E4E
U+9FF8	kCantonese	to4

U+9FF8	kTraditionalVariant	U+9FF3
U+9FF9	kCantonese	mat6
U+9FF9	kTraditionalVariant	U+2CD68
U+9FFA	kCantonese	mak6
U+9FFA	kTraditionalVariant	U+2A133
U+9FFB	kCantonese	bou3
U+9FFC	kCantonese	gaa1
U+241FE	kSimplifiedVariant	U+2B1E6
U+24AA4	kSimplifiedVariant	U+2A6DE
U+2696F	kSimplifiedVariant	U+300EE
U+27C0E	kCantonese	lou4
U+27C0E	kSimplifiedVariant	U+9FF2
U+29DD3	kCantonese	zi3
U+29DD3	kSimplifiedVariant	U+9FF5
U+29DD5	kCantonese	loeng4
U+29DD5	kSimplifiedVariant	U+9FF6
U+29E4E	kCantonese	jam1
U+29E4E	kSimplifiedVariant	U+9FF7
U+2A133	kSimplifiedVariant	U+9FFA
U+2A6DE	kCantonese	wing4
U+2A6DE	kTraditionalVariant	U+24AA4
U+2B1E6	kCantonese	jing4
U+2B1E6	kTraditionalVariant	U+241FE
U+2B735	kVietnamese	lâm
U+2C267	kCantonese	hon4
U+2C267	kSimplifiedVariant	U+9FF0
U+2CD68	kCantonese	mat6
U+2CD68	kSimplifiedVariant	U+9FF9
U+30021	kCantonese	haan1
U+30021	kTraditionalVariant	U+81E4
U+30022	kCantonese	gaam1
U+30022	kTraditionalVariant	U+30BF2
U+3005C	kCantonese	lyun4
U+3005C	kTraditionalVariant	U+9FBB
U+300EE	kCantonese	hok6
U+300EE	kTraditionalVariant	U+2696F
U+30BF2	kCantonese	gaam1
U+30BF2	kSimplifiedVariant	U+30022

13) L2/22-014: Incorrect Radical for CJK Unified Ideograph 266B9

Comments: The group reviewed this proposal, and agreed that the *kRSKangXi* and *kRSUnicode* property values for U+266B9 should be changed from 130.6 to 74.6. However, the group felt that U+266B9 does not require 130.6 as an additional *kRSUnicode* property value.

266B9
肉 130.6
𠂔 𠂔
GHZ-32067.02 T5-2E24

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept the *kRSUnicode* property value change for U+266B9, based on L2/22-014 and as amended in Section 13 of document L2/22-022, for Unicode Version 15.0.

Action Item for Michel Suignard: Change the kRSUnicode property value for U+266B9, based on L2/22-014 and as amended in Section 13 of document L2/22-022, and provide to John Jenkins an updated UniHan15.txt data file, for Unicode Version 15.0.

Action Item for John Jenkins: Update the UniHan database to reflect the property value changes for U+266B9, based on document L2/22-014 and as amended in Section 13 of document L2/22-022, for Unicode Version 15.0.

Current:

U+266B9	kRSKangXi	130.6
U+266B9	kRSUnicode	130.6

Adjusted:

U+266B9	kRSKangXi	74.6
U+266B9	kRSUnicode	74.6

14) L2/22-027: Proposal to add derived kMandarin property values

Comments: The group reviewed this proposal, and agreed that the new *kMandarin* property values are reasonable, and should therefore be checked by the CLDR-TC.

Recommendations: The group recommends to the UTC the following disposition ➡

Action Item for Peter Edberg: Ask the CLDR-TC to check the proposed kMandarin property value additions, then report back to the UTC. See document L2/22-027 and Section 14 of document L2/22-022.

UAX #45 / U-Source Documents & Topics

The documents in this section were discussed by the group, and for convenience, the following data files, all of which are PDF attachments, include all of the recommended UAX #45 *USourceData.txt* data file changes and additions based on them:

- *USource-changes-20220114.txt*
- *USource-additions-20220114.txt*
- *USource-ExtH-20220114.txt*

These recommended UAX #45 *USourceData.txt* data file changes and additions are also shown inline, following the Recommendations. In addition, the *UTCHan2022u170-Regular.ttf* TrueType font (PDF attachment) includes the representative glyphs for adding to the UAX #45 *UTCHan.ttf* TrueType font.

15) UAX #45 Adjustments for Extension H

Although there is no formal document nor public feedback on this topic, the group was reminded by Ken Whistler prior to the meeting that Extension H needs to be properly reflected in UAX #45 and its data file, *USourceData.txt*, for Unicode Version 15.0.

First, a new Field 1 (current status) value, *H* (*Encoded in Extension H*), needs to be added to the table in [Section 2.1](#), *The Status Field*, of UAX #45 proper and to the header of its data file, *USourceData.txt*.

Second, Fields 1 and 2 of the following 161 U-Source ideographs need to reflect *H* as their current status and the proper Extension H code points, respectively (the ordering was adjusted for the characters in the range U+31456 through U+31682 per Section 05 of this document, but should be confirmed using final Extension H data):

UTC-03123	U+31350
UTC-03021	U+31379
UTC-02973	U+313BC
UTC-03108	U+313CA
UTC-03030	U+313CB
UTC-03081	U+313D5
UTC-03082	U+313E2
UTC-03036	U+313E4
UTC-03003	U+31408
UTC-03084	U+31426
UTC-03044	U+3145E
UTC-03067	U+314F0
UTC-02996	U+31580
UTC-03150	U+31583
UTC-02997	U+31587
UTC-02991	U+3158E
UTC-02983	U+3158F
UTC-03022	U+31596
UTC-02990	U+31598
UTC-03000	U+3159E
UTC-03015	U+315A2
UTC-03011	U+315A7
UTC-03024	U+315AB
UTC-03049	U+315AC
UTC-03059	U+315AD
UTC-03071	U+315B2
UTC-03034	U+315B3
UTC-03035	U+315B4
UTC-03037	U+315BE
UTC-03014	U+315C4
UTC-03063	U+315C5
UTC-03032	U+315CD
UTC-03064	U+315DA
UTC-03033	U+315E5
UTC-03046	U+315E8
UTC-03027	U+315EE
UTC-03069	U+315EF
UTC-03016	U+315F0
UTC-03085	U+31602
UTC-02992	U+31604
UTC-03060	U+31606
UTC-03066	U+31609
UTC-03055	U+3162E
UTC-03052	U+31632
UTC-02981	U+316B1
UTC-03020	U+316B2

UTC-03017	U+316C2
UTC-03062	U+316C6
UTC-03028	U+316C7
UTC-03043	U+316CF
UTC-03042	U+316D0
UTC-03048	U+316E4
UTC-03026	U+316E6
UTC-03080	U+316E8
UTC-03045	U+316EE
UTC-03047	U+3170E
UTC-02995	U+31712
UTC-03087	U+3175F
UTC-03088	U+31761
UTC-02976	U+3176B
UTC-02977	U+3176C
UTC-02972	U+317AB
UTC-03147	U+317CA
UTC-03145	U+317D3
UTC-02978	U+317DB
UTC-00118	U+317E1
UTC-03089	U+317FA
UTC-03125	U+31816
UTC-03142	U+31839
UTC-00790	U+3184B
UTC-03004	U+3187F
UTC-03057	U+31884
UTC-03056	U+318D5
UTC-02984	U+318E4
UTC-03092	U+31903
UTC-03155	U+3195C
UTC-03053	U+31970
UTC-03019	U+31977
UTC-03002	U+31994
UTC-03068	U+31995
UTC-03117	U+31A1A
UTC-03050	U+31A49
UTC-03079	U+31A4A
UTC-01258	U+31AA3
UTC-03112	U+31AA6
UTC-03151	U+31ACB
UTC-03152	U+31B16
UTC-03156	U+31B7F
UTC-03025	U+31BDE
UTC-03007	U+31C24
UTC-02986	U+31C25
UTC-03031	U+31C27
UTC-01159	U+31C2D
UTC-03061	U+31C2E
UTC-03041	U+31C33
UTC-03074	U+31C50
UTC-03094	U+31CBD
UTC-03095	U+31CD2
UTC-03012	U+31CED
UTC-02982	U+31CFE
UTC-03029	U+31D07

UTC-03073	U+31D10
UTC-02999	U+31D29
UTC-03070	U+31D2A
UTC-03146	U+31D3C
UTC-03122	U+31D54
UTC-02974	U+31DC8
UTC-03096	U+31DDC
UTC-03008	U+31E11
UTC-02998	U+31E16
UTC-02985	U+31E1B
UTC-02988	U+31E25
UTC-03023	U+31E27
UTC-03013	U+31E28
UTC-03157	U+31E2D
UTC-02979	U+31E4D
UTC-01162	U+31E7C
UTC-03005	U+31E99
UTC-03039	U+31EB2
UTC-03038	U+31EB3
UTC-03018	U+31ECB
UTC-02980	U+31F66
UTC-03040	U+31F6A
UTC-03097	U+31F7A
UTC-03099	U+31F9F
UTC-03100	U+31FA0
UTC-03113	U+31FF4
UTC-02975	U+32003
UTC-03051	U+3200A
UTC-03001	U+3201D
UTC-03143	U+32043
UTC-03144	U+32048
UTC-03072	U+320BA
UTC-00315	U+320C7
UTC-03153	U+320CD
UTC-03137	U+320FD
UTC-03139	U+320FF
UTC-02970	U+32100
UTC-03140	U+32109
UTC-03107	U+32111
UTC-02987	U+32117
UTC-03093	U+32149
UTC-03010	U+3214C
UTC-03006	U+32152
UTC-03077	U+32199
UTC-03131	U+321AB
UTC-01307	U+321E0
UTC-03065	U+3223C
UTC-03078	U+3224D
UTC-03129	U+322B5
UTC-03075	U+322B9
UTC-03101	U+322FC
UTC-03114	U+32302
UTC-03148	U+32303
UTC-03141	U+3230F
UTC-03102	U+32322

UTC-01311 U+32344
UTC-02989 U+32353
UTC-03054 U+3237C
UTC-03158 U+3238E
UTC-03124 U+323A0

Lastly, a new PRI for a proposed update of UAX #45 needs to be opened, with a closing date of 2022-04-10.

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Authorize a proposed update of UAX #45 for Unicode 15.0 to update the table in Section 2.1 to reflect “H” (“Encoded in Extension H”) as a new status.

Action Item for John Jenkins, Editorial Committee: Update the table in Section 2.1 of UAX #45 to reflect “H” (“Encoded in Extension H”) as a new status, based on Section 15 of document L2/22-022, for Unicode Version 15.0.

Action Item for John Jenkins: Update the UAX #45 data file, USourceData.txt, to reflect “H” (“Encoded in Extension H”) as a new status in its header, and update the records of 161 U-Source ideographs to reflect “H” in Field 1 and the correspond Extension H code point in Field 2, based on Section 15 of document L2/22-022, for Unicode Version 15.0.

Action Item for Rick McGowan: Post the PRI for the proposed update of UAX #45, to close 2022-04-10.

16) L2/22-009: Proposal to Add to UAX #45 Three Ideographs with Radical 見

Comments: The group reviewed this proposal, and agreed that adding three new ideographs to UAX #45 with a status value of N is reasonable. The three ideographs can therefore be considered for inclusion in the UTC’s submission for the next IRG working set.

𪛗 𪛘 𪛙

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept three new U-Source ideographs as **UTC-03252** through **UTC-03254** with a UAX #45 status value of N, based on document L2/22-009 and Section 16 of document L2/22-022, for Unicode Version 15.0.

Action Item for John Jenkins: Add two new records to USourceData.txt and their representative glyphs to USourceGlyphs.pdf, based on document L2/22-009 and Section 16 of document L2/22-022, for Unicode Version 15.0.

Add:

UTC-03252;N;;147.14;;𪛗見𪛗;UTCDoc L2/22-009 1;;21;2
UTC-03253;N;;147.7;;𪛘見見;UTCDoc L2/22-009 2;kSimplifiedVariant U+32053;14;2
UTC-03254;N;;147.21;;𪛙見見見見見見;UTCDoc L2/22-009 3;;28;2

17) L2/22-011: UAX #45 Data Issues in Unicode 14.0

Comments: The group reviewed this proposal, and agreed that all of the suggested UAX #45 data file changes are reasonable. Additionally, the group agreed that the representative glyph for UTC-00443 𐀬 should be modified so that it is identical to that of UTC-00345 𐀬.

Recommendations: The group recommends to the UTC the following disposition ➡

Consensus: Accept the proposed UAX #45 data file and representative glyph changes, based on document L2/22-011 and as amended in Section 17 of document L2/22-022, for Unicode Version 15.0.

Action Item for John Jenkins: Update the UAX #45 data file, USourceData.txt, and modify the representative glyph for UTC-00443 to be identical to that of UTC-00345, based on document L2/22-011 and as amended in Section 17 of document L2/22-022, for Unicode Version 15.0.

Current:

```
UTC-00006;E;U+2B825;1.5;;𠂔乙丙;kMatthews 4150;;6;3  
UTC-00116;U;U+7AB0;116.10;;𦰩穴𦰪肉缶;kCowles 1390;;15;3  
UTC-00181;C;U+2B08E;118.9;;𦰫簾大;kSBGY 278.02;;15;3  
UTC-00193;B;U+272A5;142.6;;𧈀虫戍;kSBGY 474.14;;12;1  
UTC-00220;B;U+20957;22.11;;𨺲匚𨺳食人;kSBGY 248.24;;13;3  
UTC-00443;N;;74.8;;𣎵月枕;kCowles 129;kCantonese zam2*kDefinition callous;12;1  
UTC-00876;N;;118.6;;𬇕竹𬇖卩女;Adobe-CNS1 C+14165;;12;2  
UTC-00877;N;;188.17;;𥄑骨藏;Adobe-CNS1 C+16082;;26 27;1  
UTC-00879;N;;157.8;;𢆶足奉;Adobe-CNS1 C+16120;;15;  
UTC-00880;N;;120.13;;𡗓糸𡗒厂𡗓田兀土;Adobe-CNS1 C+16137;;19;  
UTC-00881;N;;191.13;;𡗏鬥黽;Adobe-CNS1 C+16155;;23;  
UTC-00882;N;;30.9;;𡗐口思;Adobe-CNS1 C+16196;;12;  
UTC-00883;N;;140.12;;𡗑艹𡗒白辛;Adobe-CNS1 C+16212;;16;  
UTC-00884;N;;5.5;;𡗓承乙;Adobe-CNS1 C+16219;;6;  
UTC-00886;N;;27.14;;𡗒广𡗓催月;Adobe-CNS1 C+16237;;16;  
UTC-00887;N;;85.13;;𡗓彡𡗓虎𡗓八目;Adobe-CNS1 C+16248;;16;  
UTC-00888;N;;122.9;;𡗓乛卦;Adobe-CNS1 C+16252;;14;  
UTC-00889;N;;167.10;;𡗓金祖;Adobe-CNS1 C+16257;;18;  
UTC-00890;N;;145.3;;𡗓衤丸;Adobe-CNS1 C+16290;;8;  
UTC-00891;N;;61.4;;𡗓毛心;Adobe-CNS1 C+16295;;8;  
UTC-00893;N;;45.4;;𡗓屯千;Adobe-CNS1 C+16318;;7;  
UTC-00894;N;;30.10;;𡗓口敕;Adobe-CNS1 C+16333;;13;  
UTC-00895;N;;112.14;;𡗓回白爻𡗓一八石;Adobe-CNS1 C+16340;;19;  
UTC-00896;N;;119.10;;𡗓米曷;Adobe-CNS1 C+16342;;16;  
UTC-00897;N;;5.11;;𡗓𡗓艸口冊丿;Adobe-CNS1 C+16376;;12;  
UTC-00898;N;;61.12;;𡗓卜焦;Adobe-CNS1 C+16390;;15;  
UTC-00899;N;;72.9;;𡗓日𡗓白龜𡗓丶、十;Adobe-CNS1 C+16571;;13;  
UTC-00900;N;;68.8;;𡗓斗斗斗;Adobe-CNS1 C+16806;;12;  
UTC-00910;N;;30.15;;𡗓口劈;Adobe-CNS1 C+17329;;18;  
UTC-00911;N;;32.3;;𡗓土干;Adobe-CNS1 C+17331;;6;  
UTC-00912;N;;44.6;;𡗓尸他;Adobe-CNS1 C+17334;;9;  
UTC-00913;N;;104.22;;𡗓疒𡗓彡隹女;Adobe-CNS1 C+17362;;9;
```

Adjusted:

UTC-00006;E;U+2B825;1.5;;𐤀𐤋𐤍; kMatthews 4150;;6;3
UTC-00116;U;U+7AB0;116.10;;𐤁𐤕𐤍; kCowles 1390;;15;3
UTC-00181;C;U+2B08E;118.9;;𐤁𐤕𐤍; kSBGY 278.02;;15;3

UTC-00193;B;U+272A5;142.6;;𪛗虫戌;kSBGY 474.14;;12;1
UTC-00220;B;U+20957;22.11;;𪛗食人;kSBGY 248.24;;13;3
UTC-00443;**UTC-00345**;130.8;;𪛗月枕;kCowles 129;kCantonese zam2*kDefinition cal-
lous;12;1
UTC-00876;**X**;118.6;;𪛗竹𪛗卩女;Adobe-CNS1 C+14165;;12;2
UTC-00877;**X**;188.17;;𪛗骨藏;Adobe-CNS1 C+16082;;26 27;1
UTC-00879;**X**;157.8;;𪛗足奉;Adobe-CNS1 C+16120;;15;
UTC-00880;**X**;120.13;;𪛗糸𪛗厂𪛗田兀土;Adobe-CNS1 C+16137;;19;
UTC-00881;**X**;191.13;;𪛗門𪛗;Adobe-CNS1 C+16155;;23;
UTC-00882;**X**;30.9;;𪛗口思;Adobe-CNS1 C+16196;;12;
UTC-00883;**X**;140.12;;𪛗+𪛗白辛;Adobe-CNS1 C+16212;;16;
UTC-00884;**X**;5.5;;𪛗永乙;Adobe-CNS1 C+16219;;6;
UTC-00886;**X**;27.14;;𪛗厂𪛗催月;Adobe-CNS1 C+16237;;16;
UTC-00887;**X**;85.13;;𪛗彡𪛗虎𪛗八目;Adobe-CNS1 C+16248;;16;
UTC-00888;**X**;122.9;;𪛗𪛗卦;Adobe-CNS1 C+16252;;14;
UTC-00889;**X**;167.10;;𪛗金祖;Adobe-CNS1 C+16257;;18;
UTC-00890;**X**;145.3;;𪛗𪛗丸;Adobe-CNS1 C+16290;;8;
UTC-00891;**X**;61.4;;𪛗毛心;Adobe-CNS1 C+16295;;8;
UTC-00893;**X**;45.4;;𪛗屯千;Adobe-CNS1 C+16318;;7;
UTC-00894;**X**;30.10;;𪛗口敕;Adobe-CNS1 C+16333;;13;
UTC-00895;**X**;112.14;;𪛗回白爻𪛗一八石;Adobe-CNS1 C+16340;;19;
UTC-00896;**X**;119.10;;𪛗米曷;Adobe-CNS1 C+16342;;16;
UTC-00897;**X**;5.11;;𪛗𪛗+𪛗𪛗L;Adobe-CNS1 C+16376;;12;
UTC-00898;**X**;61.12;;𪛗𪛗焦;Adobe-CNS1 C+16390;;15;
UTC-00899;**X**;72.9;;𪛗日𪛗白𪛗𪛗、十;Adobe-CNS1 C+16571;;13;
UTC-00900;**X**;68.8;;𪛗𪛗斗斗斗;Adobe-CNS1 C+16806;;12;
UTC-00910;**X**;30.15;;𪛗口劈;Adobe-CNS1 C+17329;;18;
UTC-00911;**X**;32.3;;𪛗土干;Adobe-CNS1 C+17331;;6;
UTC-00912;**X**;44.6;;𪛗尸他;Adobe-CNS1 C+17334;;9;
UTC-00913;**X**;104.22;;𪛗疒𪛗彡𪛗女;Adobe-CNS1 C+17362;;9;

UAX #50 / Vertical_Orientation Documents

The document in this section was discussed by the group.

18) **L2/21-222**: Request to revise UAX 50 for harmonization with Adobe Japan¹

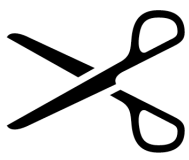
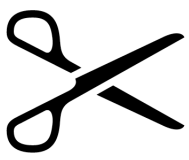
Comments: The editors of UAX #50, Ken Lunde and Koji Ishii, discussed this request to change the *Vertical_Orientation* property values for U+2016 ∥ DOUBLE VERTICAL LINE (from **U** to **R**), U+2702 ✂ BLACK SCISSORS (from **U** to **R** or **Tr**), and U+3030 ∼ WAVY DASH (from **Tr** to **R**) prior to the group meeting, and agreed that changes to their property values should be made only if there is overwhelmingly compelling evidence to do so. The group also discussed this proposal, and concluded that there was insufficient evidence for making such property value changes to these particular characters. The conclusion is that this request needs to be discussed by the UTC, and the following per-character analysis may be of help for those discussions:

U+2016 || **U**: This character is analogous to U+FF5C | whose *Vertical_Orientation* property value is **Tr**. Virtually all Japanese fonts, at least those based on [Adobe-Japan1](#) for which there are hundreds of such fonts, include an OpenType '[vert](#)' (*Vertical Alternates*) feature substitution

for U+2016 (Adobe-Japan1 CID+666), which results in Adobe-Japan1 CID+7895 (see the image below). If a property change is to be considered, **Tr** seems to be the only reasonable choice based on the fact that glyph substitution is used.

666,2	1000.00	7895,2	1000.00
			
\666		\7895	

U+2702 ✂ **U**: For virtually all Japanese fonts, at least those based on Adobe-Japan1-4 or greater for which there are hundreds of such fonts, there are four glyphs that are associated with U+2702, with Adobe-Japan1 CID+12176 being the one that is mapped from U+2702. They are Adobe-Japan1 CIDs 12175 through 12178 (see the image below). Ken Lunde claimed that treating these glyphs as though they were arrows when the OpenType 'vert' feature was updated for Adobe-Japan1-4 over 20 years ago was a mistake on his part. There is also an EVS (*Emoji Variation Sequence*) that is associated with U+2702: <2702 FE0F>.

12175,2	1000.00	12176,2	1000.00	12177,2	1000.00	12178,2	1000.00
							
\12175		\12176		\12177		\12178	

U+3030 ∞ **Tr**: Making property value changes based on a specific implementation (aka Adobe InDesign) or inconsistent browser behavior does not seem to be prudent. There is also an EVS that is associated with U+3030: <3030 FE0F>.

Recommendations: The group recommends to the UTC the following disposition ➡

Discuss the merits of changing the Vertical_Orientation property value of U+2016, U+2702, and U+3030, and record Action Items if any property value changes are to be made.

Action Item for Ken Lunde: Convey to the authors of L2/21-222 the CJK & Unihan Group comments in Section 18 of document L2/22-022, along with any feedback from the UTC.

IRG Topics

19) IRG Working Set 2021 Status

The ORT (*Online Review Tool*) for [IRG Working Set 2021 Version 2.0](#) was updated in early December 2021, and Ken Lunde finished the review of the UTC-assigned range, specifically serial numbers 02541 through 03734, on 2021-11-27. The deadline was 2022-01-14. The ORT Manager, Henry Chan, is scheduled to distribute Consolidated Comments on 2022-01-28, and IRG editors must respond to comments against their organization's submitted ideographs no later than 2022-03-04.

The IRG #58 meeting, which will be a virtual Zoom meeting, takes places from 2022-03-14 through 2022-03-18.

Recommendations: The group recommends to the UTC the following disposition ➡

No UTC action is required.

That is all.