









Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document
Title: Error Report for CJK Unified Ideographs Extensions C and E
Source: Andrew West and Ming Fan
Status: Individual Contribution
Action: For consideration by JTC1/SC2/WG2, IRG and UTC
Date: 2015-10-12

This document reports errors in the glyphs for four characters encoded in the CJK Unified Ideographs Extension C and E blocks. The source for all four characters is *Hànyǔ Fāngyán Dà Cídiǎn* 汉语方言大词典 [Great Dictionary of Chinese Dialects] (Beijing, 1999), and images from this dictionary given below show the correct glyph forms for these characters. The glyphs for these four characters should to be corrected, and in two cases the stroke count needs to be adjusted.

Code Point	Code Chart Image	Correct Glyph	Radical/ Strokes	Source Reference	Notes
2B5AA	2B5AA 页 181'.3  GZFY-00924		181'.3	<i>Hànyǔ Fāngyán Dà Cídiǎn</i> p. 4185	Simplified form of U+29454 頔 The code chart glyph has an incorrect left hand side component.
2B9D7	2B9D7 匚 22.11  GZFY-00836		22.12	<i>Hànyǔ Fāngyán Dà Cídiǎn</i> p. 2505	The code chart glyph is missing a horizontal stroke, and the stroke count is incorrect.
2BF4A	2BF4A 手 64.8  GZFY-00688		64.8	<i>Hànyǔ Fāngyán Dà Cídiǎn</i> p. 5386	Simplified form of U+3A6D 擗 The code chart glyph has the wrong top right component. 擗 is a different, non-cognate character (variant of U+632C 擗) that is currently unencoded. These two characters (擗 and 擗) are given separate entries in <i>Zhōnghuá Zìhǎi</i> 中华字海 (Beijing, 2000), and cannot be unified.
2C88D	2C88D 见 147'.7  GZFY-00673		147'.6	<i>Hànyǔ Fāngyán Dà Cídiǎn</i> p. 4605	Simplified form of U+27835 覡 The code chart glyph mistakenly has 束 instead of 束 on the left hand side, and the stroke count is incorrect.

秃

【秃】 kú ㄎㄨˊ ㄉㄨˊ 〈形〉秃。客话。罗翹云《客方言·释言》：“秃曰～。《说文》：‘～，秃也。’今语谓毫颖既尽之笔曰～笔，锋铄既尽之刀曰～刀。”

奩

【奩、奩、奩、奩】 lián ㄌㄧㄢˊ

【奩妆】 ①〈名〉嫁妆。赣语。福建建宁 [liam²¹³ tsɔŋ³⁴].

②〈名〉添箱。闽语。福建顺昌 [liē³¹ tsɔ̃⁴⁴].

【奩房】 〈名〉陪嫁的家具。冀鲁官话。山东寿光 [iǎ̃⁴³⁻³⁴ faŋ⁰].

揅

【揅】 xiǎo ㄒㄩㄥˇ ㄒㄩㄥˇ ①〈动〉捣乱。吴语。上海 [ziɔ²³]

依专门脱我瞎～你老是跟我乱捣乱。②〈动〉搅拌。湘语。湖南西部。

揅

同“揅”。见《龙龕》。

揆 “擗”的类推简化字。

Hànyǔ Fāngyán Dà Cídiǎn p. 4605

覷 【覷】 cī ㄘ ①〈动〉看。吴语。浙江温州 [ts'ɿ⁴²] 眼镜带（戴）起～就明大显。王伦《新方言杂记》：“《玉篇》：‘～，见也，盗视貌也。’今温州谓看曰～。～音如刺。”②〈动〉窥视；偷看。⊖吴语。江苏常州 [ts'ɿ⁴⁴] 矮蹲勒监头～啊～格老别在这儿探头探脑地偷看。⊖赣语。湖南浏阳 [ts'ɿ³³] 往门缝里～。③〈动〉凝视；注视。赣语。湖南浏阳 [ts'ɿ³³] 莫总～叮它，耽误事。安徽岳西 [ts'ɿ²¹]～着人。

Title: Feedback to IRG 2108; Unification between 束 versus 束

Source: suzuki toshiya

Type: individual contribution

Action: review in IRG#45 meeting

During the review of IRG N2108, I expressed a concern that the correction for this codepoint

2C88D	2C88D 见 147.7 覩 GZFY-00673	覩	147.6	Hànyǔ Fāngyán Dà Cídiǎn p. 4605	Simplified form of U+27835 覩 The code chart glyph mistakenly has 束 instead of 束 on the left hand side, and the stroke count is incorrect.
-------	-------------------------------------	---	-------	------------------------------------	--

may cause non-unifiable change of the representative glyphs. I checked the existing disunification examples out of URO (we should care for source separation rule) or Ext B (many mistakenly disunification). Yet I have no time to identify the semantics of them (because most are T-single sources), but the some cognate-but-disunified pairs would be:

850C 艸 140.11 𦵑 𦵒 𦵓 𦵔 𦵕 G0-5D78 HB2-E75C T2-5345 J1-5930 K2-5945	2C75C 艸 140.10 𦵑 TD-546C
41FF 竹 118.7 𦵑 JA-2550	7B56 竹 118.6 𦵑 𦵒 𦵓 𦵔 𦵕 𦵖 G0-325F HB1-B5A6 T1-612C J0-3A76 K0-737E V1-6174
2B0A7 竹 118.14 𦵑 TE-3C5B	2B0A1 竹 118.12 𦵑 TE-2F22
2C0D4 木 75.9 𦵑 TD-3C45	2C0F7 木 75.10 𦵑 TD-4D2D 68D7 木 75.8 𦵑 𦵒 𦵓 𦵔 𦵕 𦵖 G1-5466 HB1-B4C7 T1-5F6C J0-5C27 K0-704A V1-5A2C

Nothing to say, there are many non-cognate pairs showing this component contrast.

Also I raise another concern that CJK F is not reviewed with an assumption “束 and 束 might be unifiable”, so marking this shape difference as unifiable can cause another revisiting of CJK F.

(end of document)