

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/IRG N 1938

Date: 2013-05-17

Doc. Type:	Member body contribution
Title:	Possibly Unifiable/Duplicate Ext-E Ideographs
Source:	Kawabata, Taichi
Status:	Input to IRG
Action:	for discussion
Distribution:	IRG Members and Ideographic Experts
No. of pages:	2
Appendixes	1
Medium:	Electronic

Followings are the possibly duplicate/unifiable ideograph pairs in Extension-E.
Confirmation at IRG meeting is requested. If they are confirmed to be duplicates, I would like them to be reported to next WG2.

<p>2C7E1</p> <p>艸 140.18</p> <p style="text-align: center;">𪔐</p> <p style="text-align: right;">UTC-00839</p>	<p>𪔐</p> <p>UTC-00839</p>	<p>9FC0 (same id)</p>
---	---------------------------	-----------------------

<p>2B934</p> <p>八 12.11</p> <p style="text-align: center;">𪔑</p> <p style="text-align: right;">GKX-1393.07</p>	<p>𪔑</p> <p>GKX-0504.05</p>	<p>23344</p>
--	-----------------------------	--------------

<p>2BBCF</p> <p>夂 35.9</p> <p style="text-align: center;">𪔒</p> <p style="text-align: right;">GKX-0369.24</p>	<p>𪔒</p> <p>GKX-0152.02</p>	<p>208F8</p>
---	-----------------------------	--------------

<p>2C163</p> <p>欠 76.5</p> <p style="text-align: center;">𪔓</p> <p style="text-align: right;">GKX-0567.03</p>	<p>𪔓</p> <p>GKX-0567.03</p>	<p>2389F (same id)</p>
---	-----------------------------	------------------------

2C156
木 75.19
檫
GKX-0563.27

檫
GHZ-21324.13
23839

2C1BF
比 81.10
臭
GKX-0591.13

臭
GHZ-10352.10
23B0E

CJK E G_K glyph	Evidence for CJK E	Reference in Evidence	Similar Existing character	G-Evidence of Existing character	Scanned images from references in evidence	
<div>2B934 八 12.11</div> <div>𦵏</div> <div>GKX-1393.07</div>	章古文莫虞莫	唐韻、集韻、韻會、說文、広韻、他	𦵏U+23344 GKX-0504.05 T6-5346	𦵏玉篇古文韋字註詳部首	<div>𦵏(大広益会玉篇・宮内庁本)</div> <div>𦵏(大広益会玉篇・澤存堂本)</div> <div>𦵏(大広益会玉篇・元刊本)</div>	康熙字典(檢索本) index includes GKX-0504.05, but does not include GKX-1393.07 (?). <div>𦵏504</div> 康熙字典(標点整理本) uses same glyph (KX-0504.05) for both of KX-0504.05 and KX-1393.07 position. 《玉篇》: 古文韋字。註詳部首。 𦵏 韋韋古文莫虞莫wéixī 御定康熙字典 uses different glyphs for KX-0504.05 and KX-1393.07. However, glyph at KX-0504.05 seems to be wrong <div>𦵏章古文莫虞莫</div>
<div>2BBCE 父 35.9</div> <div>復</div> <div>GKX-0369.24</div>	復古文復復	他唐韻、集韻、韻會、正韻、說文、正字通、書彞典、書說命、禮曲禮、周礼天官、諸葛亮出師表	復U+208F8 GKX-0152.02 T6-5073	復集韻徻古作復今文作復註見丩部九畫	<div>復上同(大広益会玉篇・宮内庁本)</div> <div>復上同(元刊本)</div> <div>復上同(澤存堂本)</div> <div>復徻復古文重也(集韻・北京図書館本)</div> <div>復說同文徻(正字通)</div>	康熙字典(檢索本) index includes GKX-0152.02, but does not include GKX-0369.24 (?) at 12-strokes ノ乙 section. <div>復152</div> 復古文復復 復徻復集韻 康熙字典(標点整理本) uses same glyph (KX-0504.05) for both of KX-0504.05 and KX-1393.07 position. 11 復徻《集韻》: 徻, 古作復。《玉篇》: 同徻。今文作復。 復徻古文復復fùㄈㄨˋ《唐韻》《集韻》《韻會》

2C163
欠 76.5
𑀓𑀸𑀓𑀾𑀢𑀺
GX-0567.03

目次

吹

玉篇戶甘切集韻胡甘切丛音
酣或也與澗同揚子方言沅湘

人凡言或如此者曰澠
如是互見水部澠字註

玉篇、集韻、揚子方言

𦵏

U+2389F

GHZ-32136.130

KP1-4E86

T6-3179

(IRGKX 567.030)

(see right
column)

敢

𣎵 (一) hán 《集韻》胡甘切，平談匣。
同“澗”。《集韻·談韻》：“澗，或作𣎵。”

(二) xián 《字彙》許咸切。
气起貌。《字彙·欠部》：“𤑔，氣起貌。”

(漢語大字典)

敢
 戸甘切
 湘人言也
 (大広益会玉篇・宮内庁本)

敢

元刊本

散

湘人言也（澤存堂本）

海

敢

方言渚或也沅澧之間凡言或如此者
曰渚如是一曰渚胡不定也或作𡵚

(集韻·北京圖書館本)

耿

胡切音含或也凡言或如此者曰吹亦作渚
湖不定○又許咸切賦平聲氣起貌

(字彙)

中華大字典 had already corrected

KX shape to HYD shape. Most following dictionaries did not use KX shape.

〔耽〕胡甘切音酣單韻
沅湘人言也。見〔玉篇〕。〔按廣韻
廿三談作郎。方言作澗。集韻廿三
談云澗或作一詳水部澗字。〕

あるひは。湖(一七五八)に同じ。(玉篇)
 𪗇、沭湖人言也(集韻) 湖、方言、湖、或也、
 沭、澠之間、凡言𪗇或如此者、曰𪗇如、是、
 一曰、𪗇湖不定也、或作𪗇。

欽

16385 (中文大辞典)

康熙字典(檢索本) uses HYD shape,
too.

𪔐 567

康熙字典(上海辭書出版社) uses KX shape.

𣶒 hān 尸弓玉篇戶甘切。《集韻》胡甘切。並音酣。或也。與澗同。揚子《方言》：沅、湘人凡言或如此者曰澗如是。互見水部澗字註。

御定康熙字典 uses KX shape

玉篇戶甘切集韻
酬或也與澍同揭

<div> <div>2C1BF</div> <div>比 81.10</div> <div>𪚩</div> <div>GKX-0591.13</div> </div>	<div> <div>𪚩</div> <div>唐韻集</div> <div>韻林古</div> <div>似徃徃</div> <div>一日似狸</div> </div>	<div> <div>唐韻、集韻、說文</div> </div>	<div> <div>𪚩</div> <div>U+23B0E</div> <div>GHZ-10352.10</div> <div>T4-3E26</div> </div>	<div> <div>(see right column)</div> </div>	<div> <div> <div>𪚩</div> <div>𪚩</div> <div>說文・𪚩部</div> </div> <div> <div>《說文・𪚩部》：“𪚩，獸也，似狔狔。从𪚩，夊聲。”段注本作“狔狔”。吳大澂《古籀補》：“許書‘𪚩’字疑即石鼓‘𪚩’字，後人傳寫之誤。”</div> <div>jué《廣韻》古穴切，入屑見。月部。</div> </div> <div> <div>兽名。象猩猩。《說文・𪚩部》：“𪚩，獸也，似狔狔。”段玉裁注：“《曲禮》曰：‘狔狔能言，不離禽獸。’諸家說：狔狔如狗，聲如小兒號。其字亦作猩猩。”一說為“𪚩”的訛字。《廣雅・釋獸》：“𪚩，狹也。”《玉篇・𪚩部》：“𪚩，獸，似狸。”清鈕樹玉《說文解字校錄》：“《玉篇》‘𪚩’上有‘𪚩’，古穴切。獸，似狸。‘𪚩’下有‘𪚩’，生翼切。獸，似狸。疑實一字。石鼓文有‘𪚩’，疑《說文》是‘𪚩’。”</div> </div> <div> <div>(漢語大字典)</div> <div> <div>𪚩</div> <div>出獸</div> <div>玉名</div> <div>篇似</div> <div>加狸</div> <div>(唐写本唐韻)</div> </div> <div> <div>𪚩</div> <div>似狸</div> <div>名獸</div> <div>(王韻裴務齊本)</div> </div> <div> <div>𪚩</div> <div>狸獸</div> <div>似</div> <div>(王韻完本)</div> </div> <div> <div>𪚩</div> <div>似狔狔</div> <div>說文獸也</div> <div>(集韻·北京圖書館本)</div> </div> <div> <div> <div>𪚩</div> <div>𪚩</div> <div>許書𪚩字疑即石鼓</div> <div>𪚩字後人傳寫之誤</div> <div>𪚩</div> <div>井季龜圭</div> </div> <div>(說文古籀補)</div> </div> </div> </div>	<div> <div>KX shape was already known before KX, 正字通 described it as broken shape. However the right shape was not given.</div> <div> <div>𪚩</div> <div>也似狔狔</div> <div>舊音厥</div> <div>說文獸</div> <div>(正字通)</div> </div> <div> <div>中華大字典 had already corrected KX shape to HYD shape. Most following dictionaries did not use KX shape.</div> <div> <div>𪚩</div> <div>古穴切音狔狔</div> <div>獸也似狔狔見說文𪚩部</div> <div>玉篇廣韻皆云似狸存參</div> <div>按</div> <div>(中華大字典)</div> </div> <div> <div>康熙字典(檢索本) uses HYD shape, too.</div> <div> <div>𪚩</div> <div>591</div> </div> <div> <div> <div>𪚩</div> <div>唐韻集</div> <div>韻林古</div> </div> <div> <div>10</div> <div>𪚩</div> <div>jué ㄐㄨㄝˊ《唐韻》《集韻》並古穴切。音狔。《說文》：獸也。似狔狔。一曰似狸。</div> </div> </div> </div> </div></div>
--	--	----------------------------------	---	--	--	---

It seems that the editors of HanyuDaZidian had recognized as GKX-0591.13 was mistakenly designed during KX production. HZ-10360.06 is same with GKX-0591.17, but the description told as it is a broken form and the origin of the broken form is discussed;

𪚩

“𪚩”的讹字。《集韻·志韻》：“𪚩，江東呼貉為𪚩。或作𪚩。”按：《集韻》一本作“𪚩”；《玉篇》、《類篇》均作“𪚩”。

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1929

Date: 2013-05-21

Source:	John Knightley, individual (revised)
Meeting:	IRG#40, Hong Kong SAR, China
Title:	Response to IRGN 1938
Actions required:	FYI
Distribution:	IRG
Medium:	Electronic
Pages:	1

IRGN 1938 Possibly Unifiable/Duplicate Ext-E Ideographs notes six possible duplicates. Ext E 00228 VO4-4071 𠂔 𠂔 天 appears to be unifiable with U+4EF8 𠂔 𠂔 天.

The character U+4EF8 𠂔, IDS 𠂔 天 has two readings in Mandarin Chinese yǎo meaning weak and an older usage fó when it has the same meaning as 佛 namely Buddha (<http://www.zdic.net/zd/zi/ZdicE4ZdicBBZdicB8.htm>).

VO4-4071 𠂔 𠂔 𠂔 𠂔 𠂔 is found in both in Vietnam in Chu Nom sources and the neighbouring Guangxi in Sawndip sources though the reading changes to reflect local pronunciation the meaning is still that of the older Chinese they mean Buddha, therefore U+4EF8 𠂔 and VO4-4071 𠂔 that is they are cognates. U+4EF8 and VO4-4071 have the same abstract shape. Since they are cognates and have the same abstract shape they should be unified.

Example usage:-

“Sawndip Sawdenj” page 37

𠂔 𠂔 𠂔 bwd[pwt⁸] 佛。

“Sawndip Sawdenj” page 150

𠂔 (𠂔、勿)
faed [fat⁸] 佛：～
老。faedlaux. 佛爷。

in the index page 11 these are shown as one character:-

伏

37

878

150

ISO/IEC JTC1/SC2/WG2/IRG N1938 Feedback

Universal Multiple-Octet Coded Character Set International Organization for Standardization

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG

Title: Response to IRG N1938

Source: Dr. Ken Lunde (小林 劍)

Status: Individual Contribution

Action: For consideration by the IRG

Date: 2013-05-20

With regard to one of the Extension E duplicate characters that was reported by Japan in IRG N1938, specifically U+2C7E1, which shares the same prototypical glyph (𪛗) and source (UTC-00839) as U+9FC0, I can confirm that this is a genuine duplicate.

However, I would like to use this opportunity to point out that this duplicate was first reported as a duplicate, by me, as feedback to IRG N1779 (aka, Extension E Version 6.2), which was discussed during IRG37. See: <http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg37/IRGN1779USFeedback2.txt>

Also, my own notes indicate that I once again pointed out this Extension E duplicate during IRG38, which was for Version 7.1 or 7.2 (I cannot remember which, but it was in both).

Finally, I'd like to express my thanks to Japan for finding and reporting this duplicate, which is for the third time. That is all.

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N19.....

Date: 2013-05-21

Source:	Vietnam, NB
Meeting:	IRG#40, Hong Kong SAR, China
Title:	Response to IRGN 1938_John Knightley (20/05/13)
Distribution:	IRG
Medium:	Electronic
Pages:	1

As an answer to John Knightley's question about the similarity between 00228 𠂔 𠂔 天 (VO4-4071) and the existing character U+4EF8 𠂔 𠂔 天, the Vietnam NB suggests removing the concerned character from its submission list (extension E) for further investigating.

(end of document)

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N19.....

Date: 2013-05-23

Source:	Vietnam, NB
Meeting:	IRG#40, Hong Kong SAR, China
Title:	Response to IRGN 1938_John Knightley (20/05/13)
Distribution:	IRG
Medium:	Electronic
Pages:	1

The Vietnam NB confirms that these two characters: 00228 𡗗 𡗗 (VO4-4071) and U+4EF8 𡗗 𡗗 are not unifiable because they mean two different things, especially in the sense that the components 𡗗 and 𡗗 mean two completely opposite things. In this particular case, the Vietnamese character refers to Buddha, a deity and Vietnam has strong objection to the other meaning in 𡗗 (in the Chinese reference document) which has a negative intonation. Indeed, the concerned character 𡗗 𡗗 is composed of two elements "human" and "heaven", and not "human" and "die young,...".

So in order to avoid any misunderstanding, the 00228 𡗗 𡗗 (VO4-4071) should be maintained in the Ext E, under the codepoint 2B88B.

(end of document)