

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 **N3787**

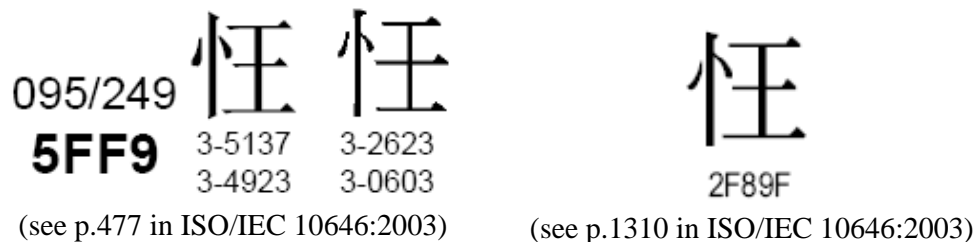
ISO/IEC JTC1/SC2/WG2 IRG N 1666

Date: 2010-3-25

Doc. Type:	Member body contribution
Title:	Request for disunifying U+2F89F from U+5FF9
Source:	TCA and China
Status:	Input to WG2 and IRG
Action:	For discussion on the meeting
Distribution:	WG2 members and IRG members
No. of pages:	6
Appendixes	None
Medium:	Electronic

1 Background

The character U+2F89F was encoded as a compatibility ideograph because it had been unified to the character U+5FF9, please refer to the figure shown below and the entry “U+2F89F” in the electronic insertion file “CJKC_SR.txt” of ISO/IEC 10646:2003. TCA and China found that both of the glyphs shown in the G-column and T-column of the code point U+5FF9 are wrong and the characters U+5FF9 and U+2F89F are non-cognate, as mentioned below.



1.1 The glyph, pronunciations and meanings of the character T3-2623

The CNS Character Database (<http://www.cns11643.gov.tw>) provides attributes of characters encoded in CNS 11643, and the Dictionary of Chinese Character Variants (<http://140.111.1.40/main.htm>) provides glyphs, pronunciations and meanings of 106,230 ideographs and the variantal relationships among them. Both of the websites are build and maintained by Taiwan’s government. It means that all information provided by those websites are offically.

By means of the CNS Character Database, the pronunciation of the character T3-2623 can be found as either “ren (ㄖㄣˋ)” or “nin (ㄋㄧㄣˊ)”, as shown following:

全字庫字型	the glyph in CNS 11643:1992		發音					
	the glyph in CNS 11643:2007		<table border="1"> <tr> <td>ㄋㄧㄣˊ</td> <td>ㄖㄣˋ</td> </tr> <tr> <td>男聲</td> <td>女聲</td> </tr> </table>		ㄋㄧㄣˊ	ㄖㄣˋ	男聲	女聲
ㄋㄧㄣˊ	ㄖㄣˋ							
男聲	女聲							
字義	筆畫	部首	簡/繁體字	倉頡				
相關詞	7	心	任	PMG				
部 件	亻, 亠, 士, 部件版本(CNS11643_2_94)							
筆順序	丶, 丨, 丶, 一, 一, 丨, 一, 筆順序版本(CNS11643_1_94)							

(see http://www.cns11643.gov.tw/AIDB/query_general_view.do?page=3&code=2623)

Most of Chinese characters are consisted of a radical part and a phonetic part. For a Chinese character having radical-phonetical structure, the radical part (namely radical) determines its class of meaning and the phonetic part (namely phonetical) determines its pronunciation and sometime more specific meaning. It's a common rule for a radical-phonetical structured Chinese character that the pronunciation of the character always similar to the pronunciation (or accient pronunciation) of its phonetical.

In the CNS 11643:1992, the character T3-2623 (任) is consisted of the radical “心” and the phonetical “王” with pronunciation “wang”. Because an inconsistency exists between pronunciations of the character T3-2623 (i.e., “ren” in offical record from CNS Character Database) and its phonetical (i.e., “wang”), the phonetical shall be wrong. That's the reason why Taiwan's MOE modified the glyph of T3-2623 to be the compositon of the radical “心” and the phonetical “壬” with pronunciation “ren” in the CNS 11643:2007.


By means of the Dictionary of Chinese Character Variants, the meaning of the character T3-2623 with pronunciation “ren” is relative to “think about”, and the meaning with pronunciation “nin” is “you (in good manners)”, as shown below.

B01102
【恁】

異
體
字

恁

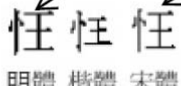
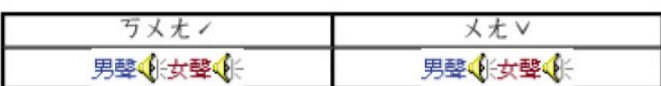
恁

音讀	(一)ㄉㄣˋ (二)ㄛˊ ㄣˋ
釋義	<p>(一)ㄉㄣˋ 思、念。文選·班固·典引：「若然受之，宜亦慙恁旅力。」宋·王安石·酬王伯虎詩：「徂年幸未暮，此意可勤恁。」</p> <p>哪。如：「恁時」。元·馬致遠·漢宮秋·第三折：「若是他不恁春風畫堂，我便官封你一字王。」</p> <p>什麼、何。初刻拍案驚奇·卷二十一：「我歇之後，有恁人在此房中安歇？」</p> <p>如此、這樣。如：「恁的」、「恁般」。宋·歐陽修·玉樓春·酒美春濃花世界詞：「已去少年無計奈，且願芳心長恁在。」明·湯顯祖·牡丹亭·第七齣：「昔氏賢文，把人禁殺，恁時節則好教鸚哥喚茶。」</p> <p>怎麼。如：「恁麼」。水滸傳·第三回：「你也須認的洒家，卻恁地教甚麼人在間壁吱吱的哭，攪俺弟兄們吃酒？」</p> <p>(二)ㄛˊ ㄣˋ 第二人稱。通「你」、「您」。董西廂·卷四：「說恁心聰，算來有分咱家共。」明·王世貞·鳴鳳記·第十四齣：「縱然恁哀鳴千狀，我此心斷易不轉。」</p> <p>作「」形時，為「恁」之異體。</p>

(see <http://140.111.1.40/yitib/frb/frb01102.htm>)

1.2 The glyph, pronunciations and meanings of the character T5-2438

By means of the CNS Character Database, the pronunciation of the character T5-2438 can be found as either “wang(ㄨㄤˋ)” or “kuang(ㄎㄨㄤˋ)”. In both of the CNS 11643:1992 and CNS 11643:2007, the character T5-2438 is consisted of the radical “心” and the phonetical “王” with pronunciation “wang”, as shown following:

全字庫字型	the glyph in CNS 11643:1992	發音
	the glyph in CNS 11643:2007	
字義	筆畫	部首
相關詞	7	倉頡
部 件	忄, 一, 土, 部件版本(CNS11643_2_94)	
筆順序	丶, 丨, 丶, 一, 一, 丨, 一, 筆順序版本(CNS11643_1_94)	

(see http://www.cns11643.gov.tw/AIDB/query_general_view.do?page=5&code=2438)

By means of the Dictionary of Chinese Character Variants, the meaning of the character T5-2438 with pronunciation “wang” is relative to “wicked”, and the meaning with pronunciation “kuang” is relative to “undisciplined” or “crazy”, as shown below.

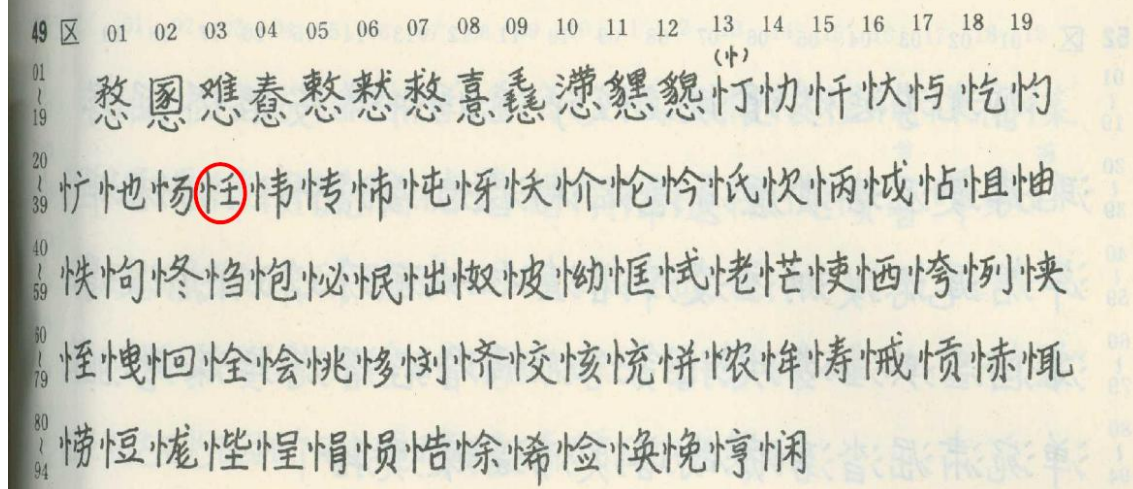
C03617 【狂】

字號	C03617	正字	【狂】 心-04-07
音讀	(一)ㄨㄤ ㄩ ㄥ (二)ㄎㄨㄤ ㄩ ㄥ		
釋義	(一)ㄨㄤ ㄩ ㄥ 邪曲。見四聲篇海·心部。 姓。見四聲篇海·心部。 (二)ㄎㄨㄤ ㄩ ㄥ 「狂」之異體。		

(see <http://140.111.1.40/yitic/frc/frc03617.htm>)

1.3 The glyph, pronunciations and meanings of G3-5137

The image below shows the glyph of U+5FF9 in G3 (GB 7589-87)



The character was selected from Page 2277 of *Hanyu Da Zidian* (漢語大字典)

任 同“恁”。《正字通·心部》：“任，同恁。”

It is considered as a variant of 恣 (U+6041) which is pronounced “ren” (meaning “thinking about”). See page 2291 of *Hanyu Da Zidian* (漢語大字典):

恣 王孫鐘
 恣 中山王鼎
 恣 說文·心部
 恣 漢印

《說文》：“恣，下齋也。从心，任聲。”段玉裁注：“未聞。按：《後漢書》班固《典引》：‘亦宜勤恣旅力。’李賢注引《說文》：‘恣，念也。’當用以訂正。”按：《食部》飪字古文作恣，古恣、飪通用。

(一) rēn 《廣韻》如甚切，上寢日。又如林切，《集韻》加鳩切。侵部。

① 思念；念及。《廣雅·釋詁二》：“恣，思也。”《玉篇·心部》：“恣，念也。”《後漢書·班彪傳》：“若然受之，宜亦勤恣旅力，以充厥道。”李賢注引《說文》曰：“恣，念也。”宋王

While the character 任 (currently coded in U+5FF9) is pronounced “kuang” (meaning “crazy”) or “wang” (meaning “evil, irregular”), see page 2275 of *Hanyu Da Zidian* (漢語大字典):

𤝵  說文·犬部
古文

(一) **kuáng** 《廣韻》巨王切，平陽羣。陽部。

同“狂”。《說文·犬部》：“狂，新犬也。𤝵，古文从心。”

(二) **wǎng** 《改併四聲篇海》引《餘文》紆往切。

❶ 邪曲，不正。《改併四聲篇海·心部》引《餘文》：“𤝵，邪曲也。”

❷ 姓。《姓鱗·養韻》：“𤝵，出《姓苑》。”

2 Proposed Solution

2.1 Change the glyphs in G-column and T-column of U+5FF9:

In order to replace the wrong glyphs (both 𤝵) in the code point U+5FF9 with the correct glyphs, TCA and China will submit the correct glyphs of G3-5137 and T3-2623 (both 𤝵) to IRG and WG2.

2.2 Remove the glyph from U+2F89F and remove the entry “U+2F89F” from the electronic insertion file “CJKC_SR.txt” of ISO/IEC 10646:2003.

2.3 A new code point U+XXXX is assigned for encoding the character T5-2438 (𤝵), and insert a corresponding entry into the electronic insertion file “CJKU_SR.txt” of ISO/IEC 10646:2003.