

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2/IRG N1354

Date: 2007-11-12

Source:	UK
Title:	Response to IRGN1347
Status :	For discussion
Distribution:	IRG members and Ideographic Experts
Medium :	Electronic

After careful review of the proposal to encode six CJK ideographs in UCS by Japan and requests the IRG to consider the following comments by the UK. The UK notes the urgency of the request to encode these six CJK ideographs, but this however is not a reason to neglect any part of the due process to required to consider this proposal.

The evidence that the present form of the second character, ARIB ID #95, is an error is overwhelming. There are readings both On and Kun for all other five characters that show clearly they are related to existing UCS ideographs. The same five characters are all found in the proposed ideographic variation sequences of Adobe. There are no readings for ARIB ID #95, because it has not been found in any dictionary, and the only known references to the character are in earlier versions of the Adobe standard, which was in later versions corrected to U+4665 禱. It is therefore suggested that due consideration be given by ARIB to the possibility of correcting the glyph of ARIB ID #95 to U+4665.

It is also observed that in the proposal IRGN1347_wg2n3318-ARIB_CJK.pdf it is acknowledged that the characters are rare and it is not specifically requested that the characters be placed in the BMP.

Brief comments on each proposed character are given below.

1) ARIB #47

This glyph is frequent used for U+6A11 梁 as for example in Yubian page 245:-



Also in Extension D review work various member bodies including Japan comment that characters differing by U+5203 刃 vs U+5205 刃 have unifiable shape:-

G: unified
 H: unifiable with 5C83 provided they have the same meaning and
 J: unifiable/duplicate. If any objection show more evidence. (*5)

Page 12

(IRGN1353_CJK_D_V22-D_Unification.pdf page 12)

 00610 UTC00260	 206C2				H: unifiable with 206C2 provided that they have the same meaning and pronunciation J: unifiable/duplicate. If any objection show more evidence. (*5)
-----------------------	-----------	--	--	--	---

(IRGN1353_CJK_D_V22-D_Unification.pdf page 64)

And

03858 TC-4A76, KP1-4B50	梁	梁 U+06881			unified. Need T and KP confirmation.		
-------------------------------	---	--------------	--	--	--	--	--

(IRGN1353_CJK_D_V22-D_QuestionableGlyph.pdf page 4)

2) ARIB #95

It is accepted that the shape of this character is not unifiable with any encoded character and therefore if after due consideration of the facts ARIB decides not to change the glyph of this character then it should be encode as a CJK ideograph

3) ARIB #39, #67, #93 and #105

It is noted all of these four characters are

(1) found as VS18 of the corresponding base character in the ideographic variation sequences proposed by Adobe,

(2) that evidence has been presented that they are cognate

therefore asks IRG to consider carefully whether or not these have unifiable shape.