Date: 2001/12/04 IRGN881

CJK Extension C Submission Format

A. Format

KXI Index	KX radical	Stroke count	First stroke	Source	Variant
XXXX.YYZ	UCS	of phonetic	of the phonetic	Information	Ucode
	code(table			and ID	U1(,U2)
	76+77) +flag			SSSNNNNN	
	XXXXY				
1-8byte	9-13	14-15	16	17-24	25-35

B. Files Types

- 1. File names: C1 and C2 submission must be put into two separate files *nameC1* and *nameC2*. The content should conform to A.format in pure text file in UTF8 Basic Latin subset.
- 2. Glyph files: 128X128 pure bitmap image(SSSNNNNNN.bmp , One file for each ideograph.)

C. Explanation for item A.form.

KXI Index (**XXXX.YYZ**): where XXXX indicates the page number where the character occurs. YY indicates the position in the page. Z with value of 0 indicates that the character exists in KX and the value of 1, which we refer as virtual code, indicates that the character does not exist in the KX dictionary. In case of virtual code, the position number is then the last character of with the

Date: 2001/12/04 IRGN881

same stroke count.

KX Radical UCS code (table 76+77) +flag

The KX radical should use UCS code table 76 and 77. The **flag** with value 0 indicates that the radical is in the traditional form and 1 indicates that the radical is in simplified form.

Note: Table 75 CJK Radicals Supplement shall NOT be used.

Date: 2001/12/04 IRGN881

First Stroke of the phonetic:

The first stroke of the phonetic should be assigned a number according to the following table:

ID	Name	Name in Chinese	Glyph1	Glyph2
1	Horizontal line	横 Heng	_	/
		提 Ti		
2	Vertical line	竪 Shu		
3	Upper- Right to Lower-Left	撇 Pie	J	
4	Upper-Left to Lower-Right	點 Dian	`	
		捺 Na		
5	Corner	折 Zhe	乙	$\Gamma \rightarrow 1$
		鈎 Gou		

SSS indicates the source information which is 3 bytes long and no space.

<u>NNNNN</u> is a 5 digit number with no space or <u>-NNNN</u> where the first one is a hyphen followed by a 4 digit number with no space.

Variant Ucode U1(,U2)

This field is for non-unifiable variant information. U1 and U2 indicates the Ucode of the already coded ideographs. The field can be empty. In case of two variants, U1 should be followed by a comma, and then U2.