

ISO/IEC JTC 1/SC 2/WG 2 N _____
 ISO/IEC JTC 1/SC 2/WG 2/IRG N 1180

ISO/IEC JTC1/SC2/WG2/IRG

**Ideographic Rapporteur Group
(IRG)**

TITLE:	Proposed additions to the CJK Strokes block of the UCS
SOURCE:	IRG Rapporteur
STATUS:	Submission from the IRG to WG2
DISTRIBUTION:	Members of ISO/IEC JTC1/SC2/WG2
DATE:	2006-4-3
REFERENCE:	WG2 N2807R, N2808R, IRG N1174
ATTACHMENTS:	IRG N1181 (“Summary for Stroke submission”)

Summary

This document proposes the addition of twenty new CJK Strokes to the CJK Strokes block of the UCS.

WG2 resolution M45.34 (N2754R) expanded the scope of IRG work to include CJK Strokes. The CJK Strokes block (U+31C0..U+31EF) now contains a total of sixteen CJK Strokes (U+31C0..U+31CF), derived from HKSCS (ISO/IEC 10646:2003/Amd.1).

The IRG formed ad-hoc groups to complete the repertoire of common CJK Strokes, and finalized the repertoire in the IRG#25 meeting held in Berkeley, California, U.S.A. The twenty new CJK Strokes are proposed for code points in the range (U+31D0..U+31E3). Information on naming conventions and collation is also provided.

List of the proposed characters, character names and code positions

	31C	31D	31E				
0	一	一	乙	31C0	CJK STROKE T	31E0	CJK STROKE HXWG
				31C1	CJK STROKE WG	31E1	CJK STROKE HZZZG
1	丿	丨	㇇	31C2	CJK STROKE XG	31E2	CJK STROKE PG
				31C3	CJK STROKE BXG	31E3	CJK STROKE Q
2	㇇	ノ	㇇	31C4	CJK STROKE SW		
				31C5	CJK STROKE HZZ		
3	㇇	ノ	〇	31C6	CJK STROKE HZG		
				31C7	CJK STROKE HP		
4	㇇	、		31C8	CJK STROKE HZWG		
				31C9	CJK STROKE SZWG		
5	㇇	㇇		31CA	CJK STROKE HZT		
				31CB	CJK STROKE HZZP		
6	㇇	一		31CC	CJK STROKE HPWG		
				31CD	CJK STROKE HZW		
7	㇇	㇇		31CE	CJK STROKE HZZZ		
				31CF	CJK STROKE N		
8	㇇	㇇		31D0	CJK STROKE H		
				31D1	CJK STROKE S		
9	㇇	㇇		31D2	CJK STROKE P		
				31D3	CJK STROKE SP		
A	㇇	㇇		31D4	CJK STROKE D		
				31D5	CJK STROKE HZ		
B	㇇	㇇		31D6	CJK STROKE HG		
				31D7	CJK STROKE SZ		
C	㇇	㇇		31D8	CJK STROKE SWZ		
				31D9	CJK STROKE ST		
D	㇇	㇇		31DA	CJK STROKE SG		
				31DB	CJK STROKE PD		
E	㇇	㇇		31DC	CJK STROKE PZ		
				31DD	CJK STROKE TN		
F	㇇	㇇		31DE	CJK STROKE SZZ		
				31DF	CJK STROKE SWG		

Stroke Type Naming Conventions

Stroke types have traditionally been named using combinations of Han characters meaning ‘horizontal’, ‘vertical’, ‘bent’, ‘curved’, etc.

These characters are conveniently represented by single letters of the alphabet, which are abbreviations for the Mandarin pronunciations of these characters. The following table shows each abbreviation, in alphabetical order, followed by the corresponding Hanyu Pinyin, Han character, and approximate meaning in English. In all cases, the final alphabetic string in the character name may be converted to the sequence of Han characters for that name, simply by substituting a Latin letter in the following list for the corresponding Han character.

For example, "HZZZG" becomes 横折折折钩.

B	biǎn	扁	‘flat’	Q	quān	圈	‘circle’
D	diǎn	點	‘dot’	S	shù	豎	‘vertical’
G	gōu	鉤	‘hook’	T	tí	提	‘rising’
H	héng	橫	‘horizontal’	W	wān	彎	‘curved’
N	nà	捺	‘right-falling’	X	xié	斜	‘slant’
P	piě	撇	‘left-falling’	Z	zhé	折	‘bent’

Variants and usage examples of CJK Strokes

The following table illustrates typical variant forms of each stroke type and Han characters encoded in ISO/IEC 10646 as examples of usage.

Stroke	Name	Variants	Usage examples
一	H (橫)		一, 三, 丁, 丞, 丈, 世, 不, 上, 十, 卅, 七
 U+31C0	T (提)		冰, 淋, 治, 冶, 冽, 暴, 汜, 录, 地, 虫
	S (豎)	/	巾, 中, 串, 冫, 乍, 上, 五, 丑

丿	SG (豎鉤)		爭, 事, 求, 水
ノ	P (撇)		乂, 爻, 禾, 毛, 乏, 乖, 采, 衣, 八, 行
㇇	SP (豎撇)		乃, 月, 用, 齊, 几, 人, 班, 大
丶	D (點)	ノ	丸, 义, 永, 冰, 凡, 丹, 主, 求, 火, 刃
㇏ U+31CF	N (捺)	㇏ ㇏ ㇏	大, 人, 天, 入, 走, 边, 廷, 尺
㇏	TN (提捺)	㇏	尗, 八, 入, 廻
㇇	HZ (橫折)		口, 冂, 田, 品, 吕, 申, 甲, 圆, 𠂇
㇇ U+31C7	HP (橫撇)	㇇	又, 双, 爻, 今
→	HG (橫鉤)		疋, 了, 予, 矛, 子, 字, 疏, 写, 冫
㇇	SZ (豎折)	㇇	断, 继, 山, 互, 彙, 牙, 乐, 东

┘	SWZ (豎彎左)		肅嘯蕭簫
└ U+31C4	SW (豎彎)		區，亡，妄，四
↓	ST (豎提)		食，良，艮，很，狠，鄉，民
┘	PZ (撇折)	ㄥ	弘，公，翁
<	PD (撇點)		巡，獵，災，笛，女，𠃉
ㄣ	PG (撇鉤)		メ
ㄩ U+31C1	WG (彎鉤)		狐，嶽，貓，家，逐
ㄣ U+31C2	XG (斜鉤)		戈，弋，戰，我
ㄣ U+31C3	BXG (扁斜鉤)		心，必，沁，恣，蕊
┘ U+31C5	HZZ (橫折折)		卍
ㄣ U+31CD	HZW (橫折彎)		爻，投，朶
┘ U+31CA	HZT (橫折提)		讠，计，鳩

 U+31C6	HZG (橫折鉤)		羽, 习, 包, 勻, 葡, 用, 青, 甫, 勺, 月, 也, 乚
 U+31C8	HZWG (橫折彎鉤)		飞, 风, 瘋, 九, 几, 气, 虱
	SZZ (豎折折)	 	亞, 鼎, 𠂇, 吳, 专, 設, 廨, 奠
	SWG (豎彎鉤)		乱, 己, 巳, 巳
 U+31CE	HZZZ (橫折折折)		凸
 U+31CB	HZZP (橫折折撇)		建, 及
	HXWG (橫斜彎鉤)		乙, 𠂇, 乞
 U+31CC	HPWG (橫撇彎鉤)		阝, 队, 邮
 U+31C9	SZWG (豎折彎鉤)		号, 亏, 弓, 强, 丐, 马, 𠂇
	HZZZG (橫折折折鉤)		乃, 孕, 仍
	Q (圈)		〇, 柁, 𠂇, 𠂇

Unicode Character Properties

All proposed characters should have the following Unicode properties so that they match those of the encoded sixteen CJK Strokes.

So;0;ON;;;;;N;;;;;

Collation of CJK Strokes

Because the HKSCS strokes were added to the block before the more complete repertoire was developed, the strokes in the block are out of order, in terms of the traditional stroke classification. CJK Strokes are divided into five broad types, in Mandarin called the five 札 zhá types, since the Han character “札” exhibits each of the types: 一 | 丿 丶 乙 (héng shù piě diǎn zhé). The collation order proposed here assigns each stroke to one and only one of the five types, and further sub-classifies each stroke according to its specific features (see IRG N987). The following is an XML representation of the collation data for the proposed CJK Strokes block, in accordance with CLDR guidelines.

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE Idml SYSTEM "http://www.unicode.org/cldr/dtd/1.4/ldml.dtd">
<ldml>
  <identity>
    <version number="$Revision: 1.00 $" />
    <generation date="$Date: 2005/12/12 12:12:12 $" />
    <language type="root" /> <!-- since this is a pan-language ordering of CJK -->
  </identity>
  <collations> <!-- removing the valid sublocales until a complete list is available -->
    <collation type="IRG-Strokes" > <!-- or whatever name would be most descriptive -->
      <rules>
        <!-- 札 zhá type 1 -->
        <!-- h; t -->
        <pc> 一 丿 </pc>
        <!-- 札 zhá type 2 -->
        <!-- s; sg -->
        <pc> | 丨 </pc>
        <!-- 札 zhá type 3 -->
        <!-- p; sp -->
```

```

    <pc> 丿 </pc>
    <!-- 札 zhá type 4 -->
    <!-- d; n; tn -->
    <pc> 丶 丶 丶 </pc>
    <!-- 札 zhá type 5 -->
    <!-- hz; hp; hg; sz; swz; sw; st; pz; pd; pg; wg; xg; bxg; hzz; hzw; hzt; hzg; hxg; szz; swg; hzzz; hzzp;
    hxwg; hpwg; szwg; hzzzg; q -->
    <pc> ㇀ ㇁ ㇂ ㇃ ㇄ ㇅ ㇆ ㇇ ㇈ ㇉ ㇊ ㇋ ㇌ ㇍ ㇎ ㇏ ㇐ ㇑ ㇒ ㇓ ㇔ ㇕ ㇖ ㇗ ㇘ ㇙ ㇚ ㇛ ㇜ ㇝ ㇞ ㇟ </pc>
    </rules>
    </collation >
    </collations>
</ldml>

```

Rationale for proposed repertoire

The repertoire of strokes proposed for addition to the existing CJK Strokes block is derived from the stroke types occurring in representative forms of currently encoded UCS Ideographs. All of these proposed strokes are currently missing from the CJK Strokes block. Representative forms of some proposed CJK Strokes are similar in appearance to representative forms of some single-stroke CJK Ideographs currently encoded in various UCS blocks (URO, Ext A, Ext B, Kangxi Radicals, CJK Radicals Supplement). However, single-stroke CJK Ideographs do not have the properties of CJK Strokes, and single-stroke CJK Ideographs may in some cases exhibit a range of variation in their representative glyphs which conflates necessary distinctions for the CJK Strokes block. For example, the proposed CJK Strokes 丿 “CJK STROKE P” (U+31D2) vs. 丿 “CJK STROKE SP” (U+31D3) are conflated in the representative forms currently used for 丿 “KANGXI RADICAL SLASH” (U+2F03) and 丿 “CJK UNIFIED IDEOGRAPHS-4E3F” (U+4E3F). The proposed strokes complete the set of most commonly seen stroke types. The precise definition of the CJK Strokes block and clear differentiation of it from the blocks of CJK Ideographs and Radicals serves an essential purpose in the indexing and collation of encoded and unencoded CJK Ideographs and Radicals. It is anticipated that some very rare strokes may be proposed for future addition to this set.