

Universal Multiple-Octet Coded Character Set

International Organization for Standardization

Doc Type: Working Group Document

Title: Thirty-seven hanzi to be considered as CJK Unified Ideographs

Source: Ken Lunde, Adobe Systems Incorporated

Status: Individual Contribution

Action: For consideration by the UTC & IRG

Date: 2008-05-19

Background

The thirty-seven Adobe-CNS1-5 hanzi that are detailed in this document represent those ideographs that we feel cannot be unified with existing CJK Unified Ideographs. What is being requested of the UTC (and IRG) is that these thirty-seven ideographs be considered as future CJK Unified Ideographs, and included with other ideographs in the pipeline.

The Adobe-CNS1-5 character collection is defined in Adobe Tech Note #5080, available from Adobe Systems. It enumerates 19,088 glyphs, specifically CIDs 0 through 19087. This glyph set covers the ideographs in Big Five, CNS 11643-1992 (Planes 1 and 2 only), Hong Kong GCCS, Hong Kong SCS (-1999, -2001, and -2004), and the Hong Kong extensions set forth by DynaComware (formerly DynaLab) and Monotype.

Proposal Details

This section lists all thirty-seven hanzi, along with multiple references, to include a representative glyph image and the closest related CJK Unified Ideograph. In some cases, the corresponding glyphs are present in the current iteration of Extension D, and their references are provided. Specifically, nine are in Extension D Version 3. Our desire is that these nine ideographs remain in Extension D, and that the remaining twenty-eight ideographs be considered for Extension E. The representative glyphs are taken from the Adobe-CNS1-5 character collection, and the typeface design is Adobe Ming Std Light.

Glyph Images & Candidate Ideographs

The following table provides the glyph images for the thirty Adobe-CNS1-5 hanzi that are being considered in this document, along with the most likely candidate CJK Unified Ideograph to represent them, either as direct Extension D references, or as the most closely-related CJK Unified Ideograph. The Adobe-CNS1-5 CIDs and U-Sources are also provided.

Adobe-CNS1-5 CID/U-Source	Glyph Image	Candidate Codes	Candidate Glyph Images	Candidate CIDs
14165/UTC00876	餃	U+25C65	篠	18074
16082/UTC00877	臟			
16120/UTC00879	蹕			
16137/UTC00880	纏			
16155/UTC00881	匱	U+9B2E	匱	13585

Adobe-CNS1-5 CID/U-Source	Glyph Image	Candidate Codes	Candidate Glyph Images	Candidate CIDs
16196/UTC00882	噁			
16212/UTC00883	薛			
16219/UTC00884	𪗇			
16231/UTC00885	㗎	D V3 01017; U+54DA	㗎	18723
16237/UTC00886	膺			
16248/UTC00887	濬			
16252/UTC00888	𪗈			
16257/UTC00889	鉏	U+93BA	鉏	17399
16290/UTC00890	𪗉			
16295/UTC00891	𪗊			
16303/UTC00892	癩		癩	17362
16318/UTC00893	𪗋			
16333/UTC00894	嗽			
16340/UTC00895	礪	U+7910 U+791C	礪 礪	12144 12493
16342/UTC00896	糲			
16376/UTC00897	亂			
16390/UTC00898	慙			
16571/UTC00899	𪗌	U+66AD	𪗌	18760
16806/UTC00900	𪗍	U+2307E	𪗍	18772
17162/UTC00901	𪗎	D V3 01166		

Adobe-CNS1-5 CID/U-Source	Glyph Image	Candidate Codes	Candidate Glyph Images	Candidate CIDs
17186/UTC00902	嗜	D V3 01107		
17213/UTC00903	甯	D V3 02174		
17215/UTC00904	嚟			
17222/UTC00905	凸	D V3 00619		
17233/UTC00906	扌	D V3 03171		
17246/UTC00907	橙			
17266/UTC00908	絜	D V3 06318		
17283/UTC00909	跂	D V3 08036		
17329/UTC00910	嘒			
17331/UTC00911	圩	D V3 01383		
17334/UTC00912	屮			
17362/UTC00913	癩		癩	16303

Note that two of the hanzi are related to each other, specifically CIDs 16303 and 17362. There is a one-stroke difference between them, meaning that only one is a candidate for becoming a CJK Unified Ideograph, and the other shall be unified with it.

Ideographic Description Sequences

The following table provides the Ideographic Description Sequences (IDSes) for each of these thirty-seven Adobe-CNS1-5 hanzi, along with each component of the IDSes represented using the “U+xxxx” notation. The first column also provides radical and stroke information in parentheses, in the form of indexing radical number (1 through 214), the number of strokes in the indexing radical instance, and the number of remaining strokes, all separated by a period, along with the first-stroke data in the form a digit (1 through 5).

Adobe-CNS1-5 CID	IDS	IDS (“U+xxxx” Notation)
14165 (118.6.7, 3)	竹 卩 女	U+2FF1 U+7AF9 U+2FF0 U+961D U+6535
16082 (188.10.18, 2)	骨 藏	U+2FF0 U+9AA8 U+85CF
16120 (157.7.8, 2)	足 奉	U+2FF0 U+27FB7 U+5949
16137 (120.6.13, 3)	糸 厂 田 兀 土	U+2FF0 U+7CF9 U+2FF8 U+5382 U+2FF3 U+7530 U+5140 U+571F
16155 (191.10.13, 2)	鬥 電	U+2FF5 U+9B25 U+9EFD

Adobe-CNS1-5 CID	IDS	IDS ("U+xxxx" Notation)
16196 (30.3.9, 2)	𠄎 口 思	U+2FF0 U+53E3 U+601D
16212 (140.3.12, 1)	𠄎 卅 𠄎 白 辛	U+2FF1 U+8279 U+2FF0 U+767D U+8F9B
16219 (5.1.5, 5)	𠄎 丞 乙	U+2FF1 U+6C36 U+4E59
16231 (30.3.6, 2)	𠄎 口 朶	U+2FF0 U+53E3 U+6736
16237 (27.2.14, 1)	𠄎 厂 𠄎 隹 月	U+2FF8 U+5382 U+2FF1 U+5020 U+6708
16248 (85.3.13, 4)	𠄎 彳 𠄎 虎 𠄎 八 目	U+2FF0 U+6C35 U+2FF8 U+864D U+2FF1 U+516B U+76EE
16252 (122.5.9, 2)	𠄎 𠄎 訃	U+2FF1 U+7F52 U+8A03
16257 (167.8.10, 3)	𠄎 金 祖	U+2FF0 U+91D2 U+8893
16290 (145.5.3, 2)	𠄎 衤 丸	U+2FF0 U+8864 U+4E38
16295 (61.4.4, 3)	𠄎 毛 心	U+2FF1 U+6BDB U+5FC3
16303 (104.5.23, 2)	𠄎 疒 𠄎 彳 𠄎 山 一 黑 女	U+2FF8 U+7592 U+2FF2 U+5F73 U+2FF3 U+5C71 U+4E00 U+9ED1 U+6535
16318 (45.3.4, 3)	𠄎 屯 千	U+2FFA U+5C6F U+5343
16333 (30.3.10, 2)	𠄎 口 敕	U+2FF0 U+53E3 U+6547
16340 (112.5.14, 3)	𠄎 𠄎 白 爻 𠄎 一 八 石	U+2FF3 U+2FF4 U+2695 U+723B U+2FF1 U+4E00 U+516B U+77F3
16342 (119.6.10, 4)	𠄎 米 曷	U+2FF0 U+7C73 U+26407
16376 (5.1.11, 1)	𠄎 𠄎 卅 口 冊 乚	U+2FF0 U+2FF3 U+8279 U+53E3 U+20541 U+4E5A
16390 (61.3.12, 3)	𠄎 卜 焦	U+2FF0 U+5FC4 U+226E7
16571 (72.4.9, 2)	𠄎 日 𠄎 白 𠄎 𠄎 、 、 十	U+2FF0 U+65E5 U+2FF1 U+767D U+2FFB U+2FF0 U+4E36 U+4E36 U+5341
16806 (68.4.8, 4)	𠄎 𠄎 斗 斗 斗	U+2FF1 U+2FF0 U+6597 U+6597 U+6597
17162 (30.3.11, 2)	𠄎 口 樺	U+2FF0 U+53E3 U+687B
17186 (30.3.9, 2)	𠄎 口 𠄎 左 月	U+2FF0 U+53E3 U+2FF1 U+5DE6 U+6708
17213 (40.3.8, 2)	𠄎 宀 笛	U+2FF1 U+5B80 U+753E
17215 (30.3.15, 2)	𠄎 口 縑	U+2FF0 U+53E3 U+260A5
17222 (18.2.5, 2)	𠄎 凸 丩	U+2FF0 U+51F8 U+5202
17233 (64.3.3, 1)	𠄎 扌 小	U+2FF0 U+624C U+5C0F
17246 (75.4.16, 1)	𠄎 木 窰	U+2FF0 U+6728 U+41B8
17266 (120.6.11, 3)	𠄎 絢 糸	U+2FF1 U+7D47 U+7CF8
17283 (157.7.2, 2)	𠄎 足 乃	U+2FF0 U+27FB7 U+4E43
17329 (30.3.15, 2)	𠄎 口 劈	U+2FF0 U+53E3 U+5288
17331 (32.3.3, 1)	𠄎 土 干	U+2FF0 U+571F U+5E72
17334 (44.3.6, 5)	𠄎 尸 牠	U+2FF8 U+5C38 U+5F75
17362 (104.5.22, 2)	𠄎 疒 𠄎 彳 𠄎 黑 女	U+2FF8 U+7592 U+2FF2 U+5F73 U+21F21 U+6535

References

The primary references for these thirty-seven ideographs are the Adobe-CNS1-5 CIDs given in the first column of the first table. A large number of CID-keyed and OpenType fonts that support the glyphs necessary for Hong Kong are based on the Adobe-CNS1 character collection. The first twenty-four ideographs have their origins in Hong Kong GCCS, which was the predecessor of Hong Kong SCS which itself has gone through three iterations (1999, 2001, and 2004). Associated with Hong Kong GCCS was an informal extension that included 145 additional hanzi specified by Hong Kong's Department of

Judiciary, and the next nine ideographs have their origins there. The last four ideographs have their origins in Monotype's Hong Kong extension, which includes 471 hanzi.

The sections below provide additional references, scanned from the following sources:

- A dictionary by 楊子來 entitled 『標準中文輸入碼大字典』, published by 聚賢館文化有限公司, 1996, ISBN 962-436-287-4.
- A dictionary by 胡双宝 entitled 『简化字 繁体字 异体字辨析手』, published by 北京大学出版社, 1996, ISBN 7-301-03198-X.

Some of these references include additional information, such as related characters, readings, indexing radicals, and total number of strokes.

CID+14165 References

Hong Kong GCCS 0x8E6C

『標準中文輸入碼大字典』 Dictionary, Page 188:

竹	13	筱	HNLK	-	-	tsui	3BT	-	8E6C	-	-	*
---	----	---	------	---	---	------	-----	---	------	---	---	---

CID+16082 References

Hong Kong GCCS 0x9EC4

『標準中文輸入碼大字典』 Dictionary, Page 306:

骨	28	臟	BBTIS	-	-	-	KGC	-	9EC4	-	-	*
---	----	---	-------	---	---	---	-----	---	------	---	---	---

CID+16120 References

Hong Kong GCCS 0x9EF4

『標準中文輸入碼大字典』 Dictionary, Page 260:

足	15	躡	RMQKQ	-	-	-	HWF	-	9EF4	-	-	*
---	----	---	-------	---	---	---	-----	---	------	---	---	---

CID+16137 References

Hong Kong GCCS 0x9F4E

『標準中文輸入碼大字典』 Dictionary, Page 200:

糸	19	纏	VFMWG	-	-	-	ZEC	-	9F4E	-	-	*
---	----	---	-------	---	---	---	-----	---	------	---	---	---

CID+16155 References

Hong Kong GCCS 0x9F60

『標準中文輸入碼大字典』 Dictionary, Page 307:

鬥	23	鬮	LNXU	jiu1	kaul	kau	OAM	-	9F60	-	-	*
---	----	---	------	------	------	-----	-----	---	------	---	---	---

CID+16196 References

Hong Kong GCCS 0x9FAD

『標準中文輸入碼大字典』 Dictionary, Page 30:

口	12	噁	RWP	si1	si3	si	HOS	-	9FAD	-	-	*
---	----	---	-----	-----	-----	----	-----	---	------	---	---	---

CID+16212 References

Hong Kong GCCS 0x9FC0

『標準中文輸入碼大字典』 Dictionary, Page 240:

艸	16	薛	THAJ	xue1	sit3	sit	TWS	-	9FC0	-	-	*
---	----	---	------	------	------	-----	-----	---	------	---	---	---

CID+16219 References

Hong Kong GCCS 0x9FC8

『標準中文輸入碼大字典』 Dictionary, Page 2:

乙	6	季	NENU	cheng2	sing4	sing	IFS	-	9FC8	-	-	*
---	---	---	------	--------	-------	------	-----	---	------	---	---	---

CID+16231 References

Hong Kong GCCS 0x9FD4

『標準中文輸入碼大字典』 Dictionary, Page 27:

口	10	噪	RNHD	duo3	doh2	doh	HRD	-	9FD4	DFE1	KMSY	*
---	----	---	------	------	------	-----	-----	---	------	------	------	---

『简化字 繁体字 异体字辨析手』 Dictionary, Page 31:

噪 * D173
duǒ “噪”的异体。

CID+16237 References

Hong Kong GCCS 0x9FDA

『標準中文輸入碼大字典』 Dictionary, Page 22:

厂	16	膺	MOGB	ying1	zing1	ying	9LY	-	9FDA	-	-	*
---	----	---	------	-------	-------	------	-----	---	------	---	---	---

CID+16248 References

Hong Kong GCCS 0x9FE6

『標準中文輸入碼大字典』 Dictionary, Page 135:

水	16	澹	EYPU	jun4	zun3	chun	6KM	-	9FE6	-	-	*
---	----	---	------	------	------	------	-----	---	------	---	---	---

CID+16252 References

Hong Kong GCCS 0x9FEA

『標準中文輸入碼大字典』 Dictionary, Page 226:

网	14	罨	WLYRY	fu4	foo6	foo	KQF	-	9FEA	-	-	*
---	----	---	-------	-----	------	-----	-----	---	------	---	---	---

CID+16257 References

Hong Kong GCCS 0x9FEF

『標準中文輸入碼大字典』 Dictionary, Page 284:

金	18	鉏	CLBM	zu3	zo2	cho	5UJ	-	9FEF	-	-	*
---	----	---	------	-----	-----	-----	-----	---	------	---	---	---

CID+16290 References

Hong Kong GCCS 0xA054

『標準中文輸入碼大字典』 Dictionary, Page 221:

衣	8	袂	LKNI	wan2	zyun2	yuen	XFY	-	A054	-	-	*
---	---	---	------	------	-------	------	-----	---	------	---	---	---

CID+16295 References

Hong Kong GCCS 0xA05A

『標準中文輸入碼大字典』 Dictionary, Page 73:

心	8	悉	HUP	mao2	mo4	mo	10M	-	A05A	-	-	*
---	---	---	-----	------	-----	----	-----	---	------	---	---	---

CID+16303 References

Hong Kong GCCS 0xA062

『標準中文輸入碼大字典』 Dictionary, Page 185:

疔	28	癥	KHOK	zheng1	zing1	ching	9BJ	-	A062	-	-	*
---	----	---	------	--------	-------	-------	-----	---	------	---	---	---

CID+16318 References

Hong Kong GCCS 0xA072

『標準中文輸入碼大字典』 Dictionary, Page 70:

中	7	氈	PUHJ	tun2	tyun4	tuen	WWC	-	A072	-	-	*
---	---	---	------	------	-------	------	-----	---	------	---	---	---

CID+16333 References

Hong Kong GCCS 0xA0A5

『標準中文輸入碼大字典』 Dictionary, Page 31:

口	13	嗽	RDBK	sou4	sau3	sau	HBS	-	A0A5	-	-	*
---	----	---	------	------	------	-----	-----	---	------	---	---	---

CID+16340 References

Hong Kong GCCS 0xA0AD

『標準中文輸入碼大字典』 Dictionary, Page 175:

石	19	礮	HCMR	-	-	-	1HH	-	A0AD	-	-	*
---	----	---	------	---	---	---	-----	---	------	---	---	---

CID+16342 References

Hong Kong GCCS 0xA0AF

『標準中文輸入碼大字典』 Dictionary, Page 193:

米	16	糲	FDASH	ta4	taap3	tap	SDW	-	A0AF	-	-	*
---	----	---	-------	-----	-------	-----	-----	---	------	---	---	---

CID+16376 References

Hong Kong GCCS 0xA0D3

『標準中文輸入碼大字典』 Dictionary, Page 2:

乙	13	亂	TBU	-	-	-	TFS	-	A0D3	-	-	*
---	----	---	-----	---	---	---	-----	---	------	---	---	---

CID+16390 References

Hong Kong GCCS 0xA0E1

『標準中文輸入碼大字典』 Dictionary, Page 79:

心	15	懔	POGP	qiao2	ciu4	chiu	S0S	-	A0E1	-	-	*
---	----	---	------	-------	------	------	-----	---	------	---	---	---

CID+16571 References

Hong Kong GCCS 0xFB53

『標準中文輸入碼大字典』 Dictionary, Page 98:

日	13	嘍	AHAJ	-	-	-	JWL	-	FB53	-	-	*
---	----	---	------	---	---	---	-----	---	------	---	---	---

CID+16806 References

Hong Kong GCCS 0xFCCC

CID+17162 References

Hong Kong GCCS 0x8A42 (Department of Judiciary extension)

CID+17186 References

Hong Kong GCCS 0x8A5A (Department of Judiciary extension)

CID+17213 References

Hong Kong GCCS 0x8A75 (Department of Judiciary extension)

CID+17215 References

Hong Kong GCCS 0x8A77 (Department of Judiciary extension)

CID+17222 References

Hong Kong GCCS 0x8A7E (Department of Judiciary extension)

CID+17233 References

Hong Kong GCCS 0x8AAB (Department of Judiciary extension)

CID+17246 References

Hong Kong GCCS 0x8AB8 (Department of Judiciary extension)

CID+17266 References

Hong Kong GCCS 0x8ACC (Department of Judiciary extension)

CID+17283 References

Hong Kong GCCS 0x8ADD (Department of Judiciary extension)

CID+17329 References

Monotype Hong Kong Extension 0xFADC

CID+17331 References

Monotype Hong Kong Extension 0xFADE

CID+17334 References

Monotype Hong Kong Extension 0xFAF9

CID+17362 References

Monotype Hong Kong Extension 0xFBEE