

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2/IRG N2294

Date: 2018-01-24

Source:	Jaemin Chung
Title:	Proposal to add G source references to characters in China's telegraph codebook
Status:	Individual contribution
Action required:	To be considered by the IRG (or by the Chinese National Body)
Pages:	6

According to the article written by Professor Koichi Yasuoka (安岡 孝一) in 1997, three characters in 标准电码本 (修订本) (Standard Telegraph Codebook (revised), 1983) were not encoded in Unicode / ISO/IEC 10646.

- 「𠄎」はなぜJIS X 0221に含まれているのか —Unicode幽霊字研究— (Why is “𠄎” included in JIS X 0221?)
<http://kanji.zinbun.kyoto-u.ac.jp/~yasuoka/publications/1997-08-29.pdf>

(Note: JIS X 0221 is just a Japanese translation of ISO/IEC 10646.)

Here are the three characters, along with their respective telegraph codes:

Character image	Telegraph code
𠄎	5495
𠄎	5831
𠄎	7016

(Source of these images: <http://kanji.zinbun.kyoto-u.ac.jp/~yasuoka/kanjibukuro/china-dempo.html> (Warning: Thousands of images included))

Since then, the first one is encoded, the second one is not encoded yet, and the last one is about to be encoded. Note that none of them are given G source references.

Character image	Telegraph code	U+	Char	Current UCS Code Chart
𠄎	5495	U+9FBF	𠄎	9FBF <small>𠄎 140.8</small>  <small>UTC-00838</small>
𠄎	5831			(unencoded)
𠄎	7016	U+30FDF (tentative)	(tentative)	30FDF <small>𠄎 167.9</small>  <small>UK-02242</small>

As some G8-#### characters are actually from 标准电码本(修订本) and will likely to be given that as their genuine source (see IRG N2276), using 标准电码本(修订本) as a source for these three characters is not a problem at all.

The following three images are excerpts from 标准电码本(修订本).

(Document continued on the next page)

5309 HW F	5400 范 HZ S	5401 茄 HZ T	5402 茈 HZ U	5403 茅 HZ V	5404 茛 HZ W	5405 茛 HZ X	5406 茛 HZ Y	5407 茗 HZ Z	5408 荔 I A A	5409 茜 I A B	3560 茛 F G Y
5319 HW P	5410 艸 I A C	5411 茛 I A D	5412 茨 I A E	5413 茫 I A F	5414 茛 I A G	5415 茛 I A H	5416 茛 I A I	5417 兹 I A J	5418 茴 I A K	5419 茵 I A L	7905 茛 L S B
5329 HW Z	5420 茶 I A M	5421 茛 I A N	5422 茸 I A O	5423 茹 I A P	5424 荀 I A Q	5425 荃 I A R	5426 荀 I A S	5427 荆 I A T	5428 葶 I A U	5429 苻 I A V	1706 苴 C N Q
5339 HX J	5430 草 I A W	5431 茛 I A X	5432 荏 I A Y	5433 荐 I A Z	5434 蕙 I B A	5435 荒 I B B	5436 茛 I B C	5437 茛 I B D	5438 茛 I B E	5439 茛 I B F	3235 茛 E U L
5349 HX T	5440 荷 I B G	5441 荻 I B H	5442 茶 I B I	5443 菱 I B J	5444 茛 I B K	5445 茛 I B L	5446 莎 I B M	5447 莒 I B N	5448 莓 I B O	5449 荃 I B P	
5359 HY D	5450 莘 I B Q	5451 莞 I B R	5452 莠 I B S	5453 莢 I B T	5454 茛 I B U	5455 莆 I B V	5456 茛 I B W	5457 葶 I B X	5458 莪 I B Y	5459 莫 I B Z	
5369 HY N	5460 菟 I C A	5461 莉 I C B	5462 莽 I C C	5463 菀 I C D	5464 菁 I C E	5465 菅 I C F	5466 菇 I C G	5467 茛 I C H	5468 菊 I C I	5469 菑 I C J	0429 茗 A Q N
5379 HY X	5470 苗 I C K	5471 菖 I C L	5472 菴 I C M	5473 菘 I C N	5474 菠 I C O	5475 菜 I C P	5476 菩 I C Q	5477 菴 I C R	5478 华 I C S	5479 菰 I C T	4538 菹 G S O
5389 HZ H	5480 菱 I C U	5481 菲 I C V	5482 茛 I C W	5483 菴 I C X	5484 菹 I C Y	5485 茨 I C Z	5486 菽 I D A	5487 蕈 I D B	5488 萃 I D C	5489 茛 I D D	6397 菴 J M B
5399 HZ R	5490 菜 I D E	5491 萋 I D F	5492 萌 I D G	5493 萍 I D H	5494 菱 I D I	5495 荏 I D J	5496 菡 I D K	5497 菌 I D L	5498 菽 I D M	5499 茛 I D N	

廿(艸)

五五

虫血行

709 复 LP	5800 蜿 IPC	5801 蝥 IPD	5802 虾 IPE	5803 蝮 IPF	5804 	5805 蝶 IPH	5806 蝮 IPI	5807 蟾 IPJ	5808 蜻 IPK	5809 蝌 IPL	
719 孑 LZ	5810 	5811 螺 IPN	5812 游 IPO	5813 蝻 IPP	5814 蝴 IPQ	5815 蜗 IPR	5816 融 IPS	5817 蠓 IPT	5818 蚂 IPU	5819 * 螞 IPV	2434 蟻 DPQ
729 蚤 MJ	5820 螟 IPW	5821 萤 IPX	5822 	5823 蚝 IPZ	5824 螃 IQA	5825 蛭 IQB	5826 螯 IQC	5827 	5828 螺 IQE	5829 蝼 IQF	7911 螞 LSH
739 丘 MT	5830 螽 IQG	5831 蚩 IQH	5832 蛰 IQI	5833 蝨 IQJ	5834 螬 IQK	5835 螭 IQL	5836 	5837 螯 IQN	5838 螳 IQO	5839 蟀 IQP	
749 冷 ND	5840 蟆 IQQ	5841 蝮 IQR	5842 蟋 IQS	5843 虬 IQT	5844 螳 IQU	5845 蟒 IQV	5846 虻 IQW	5847 蟠 IQX	5848 蝉 IQY	5849 	7924 蟑 LSU
759 涇 NN	5850 	5851 蟹 IRB	5852 蚁 IRC	5853 蟾 IRD	5854 	5855 蛭 IRF	5856 羸 IRG	5857 	5858 蚤 IRI	5859 蝇 IRJ	
769 涌 NX	5860 蚤 IRK	5861 	5862 蠕 IRM	5863 螻 IRN	5864 蟻 IRO	5865 	5866 	5867 蚤 IRR	5868 蠹 IRS	5869 蛭 IRT	
779 蜡 OH	5870 蠓 IRU	5871 蠹 IRV	5872 蛊 IRW	5873 蠲 IRX	5874 	5875 蛮 IRZ	5876 虬 ISA	5877 血 ISB	5878 噬 ISC	5879 	
789 猛 OR	5880 蚶 ISE	5881 蚌 ISF	5882 	5883 众 ISH	5884 	5885 	5886 嚙 ISK	5887 行 ISL	5888 衍 ISM	5889 	
799 幅 PB	5890 	5891 街 ISP	5892 	5893 	5894 街 ISS	5895 衙 IST	5896 	5897 	5898 卫 ISW	5899 衡 ISX	

五九

70

KJG—KNB

7000	7001	7002	7003	7004	7005	7006	7007	7008	7009	
鑊	鑛		鑿		鑛		铄		鑣	
K J G	K J H	K J I	K J J	K J K	K J L	K J M	K J N	K J O	K J P	
7010	7011	7012	7013	7014	7015	7016	7017	7018	7019	
	钥	鐳	鑲		镊	铓	铈		銑	
K J Q	K J R	K J S	K J T	K J U	K J V	K J W	K J X	K J Y	K J Z	
7020	7021	7022	7023	7024	7025	7026	7027	7028	7029	
凿	锰	长	蜚	门	闩	闪	闫	闭	闾	
K K A	K K B	K K C	K K D	K K E	K K F	K K G	K K H	K K I	K K J	
7030	7031	7032	7033	7034	7035	7036	7037	7038	7039	7596
开	闩	闾	闲		间	闵	阍		闾	闾
K K K	K K L	K K M	K K N	K K O	K K P	K K Q	K K R	K K S	K K T	L G E
7040	7041	7042	7043	7044	7045	7046	7047	7048	7049	7594
	阁	阙	闾	闾	闾	闾	闾	闾	闾	闾
K K U	K K V	K K W	K K X	K K Y	K K Z	K L A	K L B	K L C	K L D	L G C
7050	7051	7052	7053	7054	7055	7056	7057	7058	7059	
阍	闾	闾	闾	闾	闾		闾		闾	
K L E	K L F	K L G	K L H	K L I	K L J	K L K	K L L	K L M	K L N	
7060	7061	7062	7063	7064	7065	7066	7067	7068	7069	
闾	闾	闾	闾	闾	闾	闾	闾	闾		
K L O	K L P	K L Q	K L R	K L S	K L T	K L U	K L V	K L W	K L X	
7070	7071	7072	7073	7074	7075	7076	7077	7078	7079	
关		闾	闾	闾		闾	闾		阜	
K L Y	K L Z	K M A	K M B	K M C	K M D	K M E	K M F	K M G	K M H	
7080	7081	7082	7083	7084	7085	7086	7087	7088	7089	
阍		阍				阍		阍	阍	
K M I	K M J	K M K	K M L	K M M	K M N	K M O	K M P	K M Q	K M R	
7090	7091	7092	7093	7094	7095	7096	7097	7098	7099	
阪	阻	阍	阿	陀	陂	附	陋	限	陌	
K M S	K M T	K M U	K M V	K M W	K M X	K M Y	K M Z	K N A	K N B	

阜(金)长门卜(阜)

七〇副页背面

七一

To China

I would like to ask the following to the Chinese NB.

Note: I will use “GT-####” for these characters (T for telegraph; #### is the four-digit code in the codebook), as I suggested in IRG N2276.

1. Add G source references and glyphs to these two characters:

U+	Char	kIRG_GSource	G glyph image
U+9FBF	萑	GT-5495	
U+30FDF (tentative)	(tentative)	GT-7016	

2. Propose to encode this character:

(Since there is only a single character to be encoded, I suggest U+9FF0 for its code point. That is, propose to encode this character as a UNC.)

G glyph image	
kIRG_GSource	GT-5831
Proposed code point	U+9FF0
IDS	𪛗 𪛗 夕 虫
kRSUnicode	142.6
kTotalStrokes	12

3. In addition, if China wants all the characters in 标准电码本(修订本) to be encoded in the BMP, then it also needs to propose to move U+30FDF to U+9FF1, with a corresponding G glyph that has GT-7016 as its G source reference. (That is, two UNC.)

(End of document)