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.

- 「啊」はなぜJIS X 0221に含まれているのか ―Unicode幽霊字研究― (Why is "啊" 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
萑	5495
鳖	5831
镂	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
萑	5495	U+9FBF	萑	9FBF 艹 140.8 萑 UTC-00838
鳖	5831		(unencoded)	
镂	7016	U+30FDF (tentative)	(tentative)	30FDF 钅 167'.9 镂 UK-02242

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)

1. Page 55 of 标准电码本（修订本），1983

	5400 范 HZS	5401 茄 HZT	5402 苤 HZU	5403 茅 HZV	5404 苃 HZW	5405 HZX	5406 茉 HZY	5407 茗 HZZ	5408 荔 IAA	5409 茜 IAB	3560 茋 FGY	廿(艸)
5309 HWF	5410 IAC	5411 IAD	5412 茨 IAE	5413 茫 IAF	5414 茭 IAG	5415 茯 IAH	5416 茱 IAI	兹 IAJ	5418 茴 IAK	5419 茵 IAL	7905 蒿 LSB	
5319 舱 HWP	5420 茶 IAM	5421 IAN	5422 茹 IAO	5423 荀 IAP	5424 荃 IAQ	5425 萄 IAR	5426 荆 IAS	5427 荢 IAT	5428 荇 IAU	5429 IAV	1706 荬 CNQ	
5329 艰 HWZ	5430 草 IAW	5431 IAX	5432 荏 IAY	5433 荐 IAZ	5434 荑 IBA	5435 荒 IBB	5436 莅 IBC	5437 荄 IBD	5438 IBE	5439 IBF	3235 荷 EUL	
5339 苊 HXJ	5440 荷 IBG	5441 荻 IBH	5442 荼 IBI	5443 莠 IBJ	5444 莅 IBK	5445 IBL	5446 莎 IBM	5447 莒 IBN	5448 莓 IBO	5449 茎 IBP		
5349 茨 HXT	5450 莘 IBQ	5451 莞 IBR	5452 莠 IBS	5453 莢 IBT	5454 觅 IBU	5455 莆 IBV	5456 莨 IBW	5457 莩 IBX	5458 莪 IBY	5459 莫 IBZ		
5359 芭 HYD	5460 菀 ICA	5461 莉 ICB	5462 莽 ICC	5463 菀 ICD	5464 菁 ICE	5465 菅 ICF	5466 菇 ICG	5467 ICH	5468 菊 ICI	5469 菑 ICJ	0429 菪 AQN	
5369 芼 HYN	5470 ICK	5471 菖 ICL	5472 菔 ICM	5473 菘 ICN	5474 菠 ICO	5475 菜 ICP	5476 菩 ICQ	5477 堇 ICR	5478 华 ICS	5479 菰 ICT	4538 蒇 GSO	
5379 苗 HYX	5480 菱 ICU	5481 菲 ICV	5482 ICW	5483 菁 ICX	5484 菹 ICY	5485 菼 ICZ	5486 菽 IDA	5487 萁 IDB	5488 萃 IDC	5489 *萜 IDD	6397 蔌 JMB	五五
5389 芦 GHZH	5490 菜 IDE	5491 萋 IDF	5492 萌 IDG	5493 萍 IDH	5494 萎 IDI	5495 萑 IDJ	5496 萏 IDK	5497 菌 IDL	5498 萩 IDM	5499 IDN		
5399 茂 QHZR												

虫血行

	5800 蜿 IPC	5801 蝥 IPD	5802 虾 IPE	5803 蝮 IPF	5804 IPG	5805 蝶 IPH	5806 蜷 IPI	5807 蟮 IPJ	5808 蜻 IPK	5809 蝌 IPL
709 ⋯ LP	5810 螺 IPM	5811 蝣 IPN	5812 蛹 IPO	5813 蝴 IPP	5814 蜗 IPQ	5815 融 IPR	5816 蠓 IPS	5817 蚂 IPT	5818 *蟎 IPU	2434 蟥 DPQ
719 ⋯ LZ	5820 蟆 IPW	5821 萤 IPX	5822 IPY	5823 蚝 IPZ	5824 螃 IQA	5825 蛳 IQB	5826 螯 IQC	5827 IQD	5828 螺 IQE	5829 蝼 IQF · 7911 螅 LSH
729 ⋯ MJ	5830 蠡 IQG	5831 鲨 IQH	5832 蛰 IQI	5833 蟊 IQJ	5834 蟥 IQK	5835 螭 IQL	5836 IQM	5837 鳌 IQN	5838 螳 IQO	5839 蟀 IQP
739 ⋯ MT	5840 蟆 IQQ	5841 蝈 IQR	5842 蟋 IQS	5843 虮 IQT	5844 蟪 IQU	5845 蟒 IQV	5846 蟛 IQW	5847 蟠 IQX	5848 蝉 IQY	5849 IQZ · 7924 蟑 LSU
749 ⋯ ND	5850 IRA	5851 蟹 IRB	5852 蚁 IRC	5853 蟾 IRD	5854 IRE	5855 蛏 IRF	5856 蠃 IRG	5857 IRH	5858 蚤 IRI	5859 蝇 IRJ
759 ⋯ NN	5860 蛋 IRK	5861 蠕 IRL	5862 蠖 IRM	5863 蟥 IRN	5864 IRO	5865 IRP	5866 蠡 IRQ	5867 蠢 IRR	5868 蛎 IRS	5869 IRT
769 ⋯ NX	5870 蟥 IRU	5871 蠹 IRV	5872 蛊 IRW	5873 蠲 IRX	5874 IRY	5875 蛮 IRZ	5876 虻 ISA	5877 血 ISB	5878 螠 ISC	5879 ISD
779 ⋯ OH	5880 衄 ISE	5881 衅 ISF	5882 ISG	5883 众 ISH	5884 ISI	5885 ISJ	5886 嘈 ISK	5887 行 ISL	5888 衍 ISM	5889 ISN
789 ⋯ OR	5890 衔 ISO	5891 ISP	5892 ISQ	5893 街 ISR	5894 衙 ISS	5895 IST	5896 ISU	5897 ISV	5898 卫 ISW	5899 衡 ISX
799 ⋯ PB										

五九

钅(金)长门阝(阜)

7000 镶 KJG	7001 镔 KJH	7002 KJI	7003 鉴 KJJ	7004 KJK	7005 钨 KJL	7006 KJM	7007 铄 KJN	7008 KJO	7009 镰 KJP
7010 KJQ	7011 钥 KJR	7012 镭 KJS	7013 镶 KJT	7014 KJU	7015 镊 KJV	**7016 铹 KJW**	7017 锣 KJX	7018 KJY	7019 銮 KJZ
7020 錾 KKA	7021 锰 KKB	7022 长 KKC	7023 蜱 KKD	7024 门 KKE	7025 闩 KKF	7026 闪 KKG	7027 闫 KKH	7028 闭 KKI	7029 闰 KKJ
7030 开 KKK	7031 闶 KKL	7032 闱 KKM	7033 闲 KKN	7034 KKO	7035 间 KKP	7036 闵 KKQ	7037 闸 KKR	7038 KKS	7039 阁 KKT	7596 阉 LGE
7040 阁 KKU	7041 阀 KKV	7042 闺 KKW	7043 闽 KKX	7044 阃 KKY	7045 阆 KKZ	7046 阅 KLA	7047 阎 KLB	7048 阅 KLC	7049 圆 KLD	7594 阅 LGC
7050 阉 KLE	7051 阐 KLF	7052 阏 KLG	7053 阄 KLH	7054 阈 KLI	7055 阊 KLJ	7056 KLK	7057 闱 KLL	7058 KLM	7059 阔 KLN
7060 阕 KLO	7061 阑 KLP	7062 阒 KLQ	7063 阗 KLR	7064 阕 KLS	7065 阊 KLT	7066 阖 KLU	7067 阙 KLV	7068 闯 KLW	7069 KLX
7070 关 KLY	7071 阛 KLZ	7072 阐 KMA	7073 阚 KMB	7074 KMC	7075 KMD	7076 阛 KME	7077 囡 KMF	7078 KMG	7079 阜 KMH
7080 阝 KMI	7081 KMJ	7082 阡 KMK	7083 KML	7084 KMM	7085 KMN	7086 阮 KMO	7087 KMP	7088 阱 KMQ	7089 防 KMR
7090 阪 KMS	7091 阻 KMT	7092 阼 KMU	7093 阿 KMV	7094 陀 KMW	7095 陂 KMX	7096 附 KMY	7097 陋 KMZ	7098 限 KNA	7099 陌 KNB

七○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	萑
U+30FDF (tentative)	(tentative)	GT-7016	镵

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	鳖
kIRG_GSource	GT-5831
Proposed code point	U+9FF0
IDS	⿱⿰⿱⿰⿰ 丬夕虫
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 UNCs.)

(End of document)

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type:	**Ideographic Rapporteur Group Document**
Title:	**Feedback on IRGN2294**
Source:	**Eiso Chan(陈永聪), Prajñavṛttā(ⴑញ᠋ꤶ឵)**
Status:	**Individual Contribution to IRG #50, Beijing, China**
Action:	**For consideration by IRG, China and UTC**
Date:	**2018-05-06**

This document just only provides some possible useful information of the three characters which are mentioned in IRGN2294 by Jaemin.

Eiso and Prajñavṛttā did a test on these three characters which their telecodes are 5495, 5831 and 7016. Prajñavṛttā went to Beijing Long-Distance Calling Building(北京长话大楼) to wire Eiso, which the requested text has been shown as below.

Telecode	5495	5831	7016	5814	5805
Character	萑	鲞	�	蝴	蝶

The staffs told Prajñavṛttā the second and the third characters, aka 鲞(UTC-03167 in UAX #45) & 锵(04531:UK-02242 in WS2015), couldn't be inputted, so they were not allowed to use for the telegrams.

On the other hand, the first character, aka 萑(U+9FBF), was treated as 萑(U+8411), please see Fig. 1. But the staffs knew they were different.

Fig. 1 U+9FBF vs U+8411
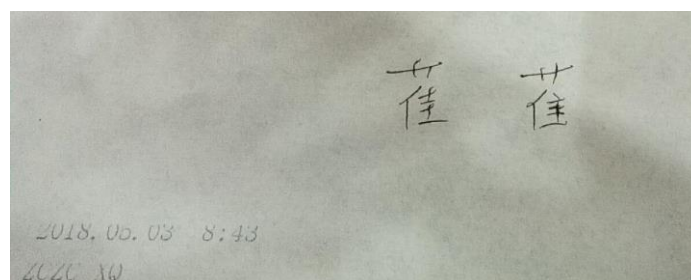
Interestingly enough, Characters 5495 in 《標准電碼本》(1952), 《新編標准電碼本》
(1972),《标准电码本（修订本 》(1983) and *Silabário Codificado de Romanização do
Cantonense* are 萑(U+9FBF), but it was shown as 萑(U+8411) in 《電碼新編》(1945), and
the current kMainlandTelegraph and kTaiwanTelegraph values for 萑(U+8411) are 5495 in
Unihan Database.

Notice that the Cantonese pronunciation for Character 5495 provided in *Silabário Codificado
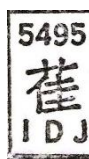de Romanização do Cantonense* is CHOI.

Fig. 2 Character 5495 in the 1952 Edition

Fig. 3 Character 5495 in the 1972 Edition in HK
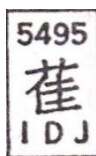
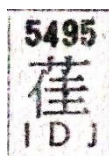*Fig. 4 Character 5495 in the 1983 Edition*

Fig. 5 Character 5495 in Silabário Codificado de Romanização do Cantonense in Macao
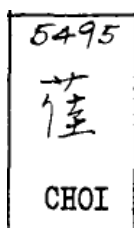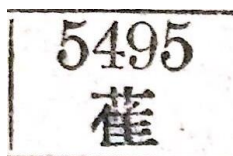

Fig. 6 Character 5495 in the 1937 Edition

Eiso received this telegram after two days. The first character, aka Character 5495, had been
changed to "0000" which means "null". We thought the staffs from China Telecom in
Guangzhou couldn't input 萑(U+9FBF) and they thought 萑(U+9FBF) and 萑(U+8411) were
different characters. Please see Fig. 7.

Fig. 7 A Telegram from Prajñavṛttā(Beijing) to Eiso(Guangzhou)

(End of Document)