Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2/IRG N2005A
Date: 2014-05-15

Source:	UTC
Title:	UTC/US Urgently-needed Character Submission
Meeting:	IRG #42
Status :	
Actions required	
Distribution:	IRG
Medium :	Electronic
Page:	
References:	

This UTC/US urgently-needed character submission consists of the following documents:

IRG N2005A:	This document
IRG N2005B:	Proposal summary form from the IRG P&P
IRG N2005C:	A table of the proposed characters and key attributes
IRG N2005D:	A Excel spreadsheet with character attributes (main data file)
IRG N2005E:	A zip archive of bitmaps
UTCHan.ttf:	A font containing glyphs for all the characters

The table (document C) includes the following attributes:

A. The U-source identifier
B. The character's glyph
C. The character's IDS (if any)
D. The KangXi radical (number form)
E. The KangXi radical (character form)
F. The number of additional strokes
G. The stroke type of the first additional stroke
H. An indication of whether the character is for traditional Chinese (TC) or simplified Chinese (SC)
I. A character of which the proposed character is a variant. In the case of SC characters, the variant is the traditional form(s) (if encoded)
J. The code point of the variant character

The Excel spreadsheet (document D) includes the following attributes:

A. The U-source identifier

B. The name of the file containing the bitmap for the glyph (always the identifier + ".bmp")

C. The KangXi radical (code point) plus an additional 0 or 1 indicating that the character is traditional Chinese or simplified Chinese, respectively

D. The number of additional strokes

E. The stroke type of the first additional stroke

F. An indication of whether the character is for traditional Chinese (0) or simplified Chinese (1)

G. The character's IDS

H. The code point of a character of which the proposed character is a variant. In the case of SC characters, the variant is the traditional form(s) (if encoded)

I. The PUA code point for the character's glyph in UTCHan.ttf font.

Note that all the data required by IRGN1920 2.2.1 and Annex G.1 is included in the Excel spreadsheet.

Rational for "Urgent Need" and Evidence

In general, the characters included here have sufficiently established current use that we feel it inappropriate to wait for encoding in Extension G, which would otherwise be the earliest opportunity.

UTC-00953 is a recent coinage for a newly-synthesized chemical element, copernicium. As such, it is required for the periodic table of the elements and other common reference works for chemistry and physics. It is the only simplified Chinese character for a currently-known element not yet encoded.

UTC-00954, UTC-00957, and UTC-00958 are comparatively common fish and plant names, found in standard reference materials in both simplified and traditional Chinese.

UTC-00959 is required by the catalogue of the Morrison collection, a major Western collection of old Chinese books and manuscripts at the SOAS at London University.

Specific evidence is below.

UTC-00953, from http://baike.baidu.com/subview/1861163/5643972.htm?fromtitle=钅哥&fromid=3513748&type=syn (fetched 15 May 2014)

鿏（化学元素） 编辑

钅哥即鿏（化学元素）。

2010年2月19日，于1996年被合成出来的第112号元素被正式命名为“Copernicium”，符号为Cn，中文译名为“鿏”。此名称是为了纪念著名天文学家哥白尼（Copernicus）而得名的。

目录

1 简述 编辑

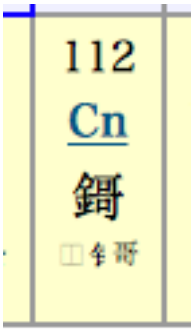
于2010年2月19日被命名。第112号化学元素，英文copernicium，缩写Cn（为纪念哥白尼），读音为“gē”。

德国重离子研究中心于1996年在粒子加速器中用锌离子轰击铅靶首次成功合成了第112号化学元素的一个原子，2002年重复相同的实验又制造出一个第112号化学元素的原子。此后，日本的一个研究机构于2004年也合成了这种元素的两个原子，从而证实德国科学家的发现。

新元素原子质量约为氢原子质量的277倍，是得到国际纯粹与应用化学联合会正式承认的最重的元素。为金属元素，具有强放射性。

鿏的简化汉字 [1]

UTC-00953, from http://kanji-database.sourceforge.net/tables/periodic.html (fetched 15 May 2014)



UTC-00953 from http://zhongxue.hujiang.com/tiku/p1389640/ (fetched 15 May 2014)

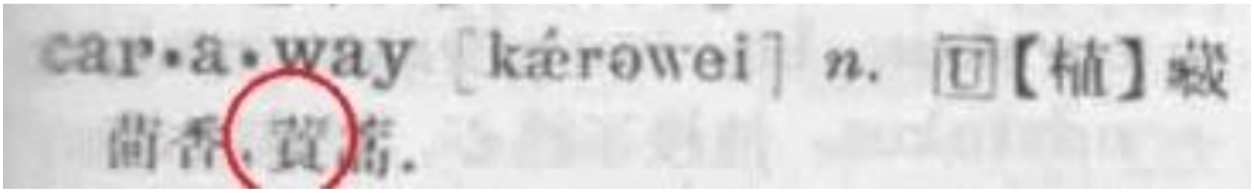
Cm	96	鋦	锔	Curium
Cn	112	鿔	钅哥	Copernicium
Co	27	鈷	钴	Cobalt

UTC-00954, from Zheng Lan-ping, Chen Xiao-yong, Yang Jun-xing, "Composition and Status of Fishes of Nanla River in Xishuangbanna, Yunnan, China"; Zoological Research 2009(3): 334-340.

附录Ⅰ 云南省南拉河鱼类名录
Appendix Ⅰ List of Fishes in Nanla River, Yunnan, China

中文名 Chinese name	拉丁名 Latin name
I. 鲤形目	CYPRINIFORMES
鲤科	Cyprinidae
1. 金线鲃	*Danio chrysotaeniata*
2. 斑尾低线鱲	*Barilius caudiocellatus*
3. 丽色低线鱲	*Barilius pulchellus*
4. 长嘴鱲	*Raiamas guttatus*
5. 高体鳑鲏*	*Rhodeus ocellatus*
6. 异斑小鲃	*Puntius ticto*
7. 中国结鱼	*Tor sinensis*
8. 云南吻孔鲃	*Poropuntius huangchunchieni*
9. 棱吻孔鲃	*Poropuntius carinatus*

UTC-00957, from Jiǎnmíng Yīng-Hàn Cídiǎn p.135.



UTC-00958, from Hénán Zhíwù Zhì 河南植物志 (1981-1998), vol.3 p.150.

15. 茴芹属 *Carum* L.

二年或多年生草本，高 30cm～80cm，根肉质。茎直立，分枝，上部 2 叉状。基生及下部的叶柄具鞘，中、上部叶有短柄或无柄；叶片二至三回羽状分裂，末回裂片线形或披针形，全缘。复伞形花序顶生或侧生，总苞和小总苞片无或 1 个～10 个，线形或披针形；小伞形花序有花 4 朵～30 朵，花两性或杂性；萼齿不显；花瓣白色或红色，倒卵形，顶端凹陷有内折小舌片；花柱基圆锥形。果实长卵形或卵形，侧扁，果棱明显；每棱槽中有油管 1 个，稀 3 个，合生面 2 条～4 条。

有 25 种～30 种，分布于欧洲、亚洲、北非及北美。我国有 4 种，2 变型。河南有 2 种。

1. 茎生叶鞘有宽白色膜质边缘。小总苞片无偶有 1 个～3 个，披针形；小伞形花序有花 5 朵～15 朵。果实长 4mm～5mm …… (1)**葛缕子** *Carum carvi* L.

1. 茎生叶鞘白色膜质边缘狭。小总苞片 5 个～8 个，线形；小伞形花序有花 15 朵～25 朵。果实长 3mm～4mm …… (2)**田葛缕** *Carum buriaticum* Turcz.

(1)**葛缕子** **葛缕子 *Carum carvi* L.** (**图 1696 7～8**)

二年生或多年生草本，高 30cm～80cm，无毛。根圆锥状，肉质。茎直立，上部分枝。基生叶

UTC-00958, from Héběi Zhíwù Zhì 河北植物志 (1986-1991), vol.2 p.256.

2. 田葛缕子 田茴芹 (图 1160)

Carum buriaticum. Tvrcz. in Bull. Soc. Nat Mosc. 11: 92. 1838. nom. nud. id. 17:713.1844. describ; 中国高等植物图鉴 2:1070. 图3869. 1972.

二年生草本，高 30—70 厘米，无毛。根纺锤形，肥厚。茎直立，基部被淡褐色基生叶残迹。基生叶具长柄，轮廓长圆状

UTC-00959, from Andrew West, *Catalogue of the Morrison Collection of Chinese Books* (SOAS, 1998), p.16.

Zengbu Man-Han zhuanzihui. 1762. RM 166.

A dictionary of Chinese characters. Compiled by Mei Yingzuo (c.1570-?) in 1615. This edition also gives the seal script form and a phonetic reading in the Manchu script for most characters. Appended are two tables by Mei Yingzuo of the possible phonetic combinations in Chinese, one arranged by rhyme and the other by alliteration. Cf. Simon & Nelson, *Manchu Books in London* II.14.

ff. 3, 3, 6, 29, 54, 51, 50, 51, 52, 53, 69, 66, 61, 71, 55, 54, 45.

RM CK494.13 166. 20 fascs. in 3 vols. 23.5 cm.

《增補滿漢篆字彙》十二集，首一卷，末一卷。（明）梅膺祚音釋，（明）廖百子（廖文英）纂輯。

清乾隆二十七年（ 1762 ）刻本，帶月樓藏版。

版式：版框 20.5 x 14.3 公分；四周單邊； 11 行 20 字，小字雙行 36 字。

扉頁：鐫“乾隆壬午〔 27 年〕春刻；梅挺生先生原本；帶月樓藏板；《增補滿漢篆書字匯彙》”。

內容：末卷收《韻法直圖》、《韻法橫圖》。

序跋：康熙壬子（ 11 年）龔鼎孳序。

Zengbu Man-Han zhuanzihui 增補滿漢篆字彙 (RM CK494.13 166): title-page.

UTC/US Urgently-needed Character Submission

Submission Contents

UTCIndex	PUAGlyph	IDS	Radical	Ra…	Strokes	FirstS…	Simplified
UTC-00953	镉	⿰钅哥	167	金	10	1	SC
UTC-00954	鲕	⿰鱼丹	195	魚	4	2	SC
UTC-00957	黃	⿱艹頁	140	艸	9	1	TC
UTC-00958	黄	⿱艹页	140	艸	9	1	SC
UTC-00959	汪	⿰氵玉	15	氵	5	1	TC

ISO/IEC JTC 1/SC 2/WG 2/IRG
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS OF CJK UNIFIED IDEOGRAPHS TO THE REPERTOIRE OF ISO/IEC 10646
Please fill in all the sections below.
Please read Principles and Procedures Document (P & P) from http://appsrv.cse.cuhk.edu.hk/~irg/irg31/IRGN1562.pdf **for guidelines and details before filling in this form.**
Please ensure you are using the latest Form from http://appsrv.cse.cuhk.edu.hk/~irg/SubmissionForm.pdf **.**
See also http://appsrv.cse.cuhk.edu.hk/~irg/UCV.html **for latest *Unifiable Calligraphic Variations*.**

A. Administrative

1. **IRG Project Code:**	UTC/US Urgently Needed Characters
2. **Title:**	UTC/US Urgently Needed Characters
3. Requester's region/country name:	USA
4. Requester type (National Body/Individual contribution):	National Body
5. Submission date:	5 May 2014
6. Requested Ideograph Type (Unified or Compatibility Ideographs)	Unified Ideographs
If Compatibility, does requester have the intention to register them as IVS (See UTS #37) with the IRG's approval? (Registration fee will not be charged if authorized by the IRG.)	N/A
7. Request Type (Normal Request or Urgently Needed)	Urgently Needed
8. Choose one of the following:	
This is a complete proposal:	Yes
(or) More information will be provided later:	

B. Technical – General

1. Number of ideographs in the proposal:	*5*
2. Glyph format of the proposed ideographs: (128x128 "bmp" files or TrueType font file)	*TrueType*
If 'bmp' files, their file names are the same as their Source IDs?	
If TrueType font, all proposed glyphs are put into BMP PUA area?	*Yes*
If TrueType font, data for Source IDs vs. character codes are provided?	*Yes*
3. Source IDs:	
Do all the proposed ideographs have a unique, proper Source ID (country/region code and less than 9 alphanumeric characters)?	Yes
4. Evidence:	
a. Do all the proposed ideographs have the separate evidence document which contains at least one scanned image of printed materials (preferably dictionaries)?	Yes
b. Do all the printed materials used for evidence provide enough information to track them by a third party (ISBN numbers, etc.)?	Yes
5. Attribute Data Format: (Excel file or CSV)	*Excel file*

C. Technical - Checklist

Understandings of the Unification Checklist	
1.Has the requester read ISO/IEC 10646 Annex S and did the requester understand the unification policy?	*Yes*
Has the requester read the “Unifiable Calligraphic Variations” (contact IRG technical editor through the Rapporteurfor the latest one) and did the requester understand the unifiable variation examples?	*Yes*
Has the requester read this P&P document and did the requester understand the 5% rule?	*Yes*
Character-Glyph Duplication Checklist(http://www.itscj.ipsj.or.jp/sc2/open/pow.htm contains all the published ones and those under ballot**)**	
Has the requester checked that any of the proposed ideographs is ***not unifiable*** with the unified or compatibility ideographs of ISO/IEC 10646?	*Yes*
If yes, which version of ISO/IEC 10646 did requester check? (e.g. 10646:2003)	10646:2011
Has the requester checked that any of the proposed ideographs is ***not unifiable*** with the ideographs in Amendments of current ISO/IEC 10646? (As of 2009, Amendment 1, 4, 5, 6and 8 have CJK ideographs.)	*N/A*
If yes, which amendments did requester check?	N/A
Has the requester checked that any of the proposed ideographs is ***not unifiable*** with the ideographs in the current IRG working sets or proposed amendments of ISO/IEC 10646? (As of 2009, PDAM 6 and PDAM 8 have CJK ideographs.)	*Yes*
If yes, which draft amendments did requester check?	Extension E
Has the requester checked that any of the proposed ideographs is ***not unifiable*** with the ideographs in the current working M-set and D-set of the IRG? (Contact IRG chief editor and technical editor through the IRG Rapporteur for the newest list)	*Yes*
If yes, which document did requester check?	IRGN 1979 (CJK Ext. F1. v2.0)
Has the requester checked that any of the proposed ideographs is ***not unifiable*** with the over-unified or mis-unified ideographs in ISO/IEC 10646? (Check Annex E of this document).	*Yes*
Has the requester checked that any of the proposed ideographs ***has similar ideograph(s)*** with the ideographs in the current standardized or working set mentioned above?	*Yes*
Has the requester checked that any of the proposed ideographs ***has variant ideograph(s)*** with the ideographs in the current standardized or working set mentioned above?	*Yes*
Attribute Data Checklist	
Do all the proposed ideographs have attribute data such as the KangXi radical code, stroke count and first stroke?	*Yes*
Are there any simplified ideographs (ideographs that are based on the policy described in簡化字總表) in the proposed ideographs?	*Yes*
If YES, does your proposal include proper simplified/traditional indication flag for each proposed ideograph in attribute data?	*Yes*
Do all the proposed ideographs have the document page number of evidence documents in attribute data?	*Yes*
Do all the proposed ideographs have the proper Ideographic Description Sequence (IDS) in attribute data?	*Yes*
If NO, how many proposed ideographs do not have the IDS?	
If the answer to question 9 or 10 is yes, do the attribute data include any information on similar/ variant ideographs for the proposed ideographs?	*Yes*