

出國報告(出國類別：出席國際會議)

出席ISO/IEC JTC1/SC2/WG2/IRG 第29次會議

服務機關：經濟部標準檢驗局

姓名職稱：莊主任秘書素琴

派赴國家：美國加州舊金山灣區

出國期間：96年11月11日至18日

報告日期：96年12月28日

摘 要

本次會議由Unicode協會主辦，共計33人出席，主要目的為審查CJK_D字集，討論急需編碼字集、專屬用字編碼區及表意文字認同規則，擬訂IRG下一階段工作方向，並確認IRG第30～32次會議舉辦地點與時間。

依IRG第28次會議（前次會議）之決議，我國已完成兩輪CJK_D字集審查工作，並依時限將審查結果提交IRG首席編輯。另有關甲骨整理工作，我國已依上次IRG決議完成《說文解字》第101～140部首之蒐集及整理，及彙整於本（96）年9月13～15日所召開古漢字工作組會議之工作成果，並提交IRG工作組。

ISO 10646系列標準可說是少數我國能夠積極參與制訂與維護的國際標準之一，後續IRG工作組將再持續進行有關漢字集的擴編、中文字的屬性及古漢字的整理等多項工作，爰我國應能長期參與該工作組活動，並積極貢獻，才能建立及維持我國在該領域的影響力。

目 次

壹、會議概要及目的-----	1
貳、會議過程-----	4
參、會議議題及決議-----	7
肆、我國應配合辦理之工作-----	11
伍、心得與建議-----	12
陸、附件-----	13

壹、會議概要及目的

一、會議時間

96年11月12日～16日。

二、會議地點

美國加州舊金山灣區Adobe公司。

三、主辦單位

Unicode協會。

四、出席人員

本次會議共計33人出席會議，分屬台灣、中國、香港、日本、南韓、越南、美國及Unicode協會，與會人員名單如下：

姓名	會員體	服務單位
陸勤(Lu, Qin)	IRG主席	香港理工大學電子計算學系
莊素琴(Chuang, Julie S. C.)	台灣(TCA)	經濟部標準檢驗局
曾士熊(Tseng, Shih-shyeng)	台灣(TCA)	中央研究院計算中心
魏林梅(Wei, Lin-mei)	台灣(TCA)	中文數位化技術推廣基金會
許其清(Hsu, Chi-ching)	台灣(TCA)	個人
黃柏盛(Hung, Fang-chuan)	台灣(TCA)	主計處電子資料處理中心
周智禾(Chou, Chih-ho)	台灣(TCA)	主計處電子資料處理中心
陳壯(Chen, Zhuang)	中國	中國電子技術標準化研究所
尹江紅(Yin, Jianghong)	中國	北京方正電子公司
張力偉(Zhang, Liwei)	中國	北京中華書局
史建橋(Shi, Jianqiao)	中國	商務印書館
鄭偉康(Cheng, Wai-hong, Peter)	香港	香港政府法定語文事務處
甄炯輝(Yan, Retarkgo)	香港	香港政府資訊科技署
關穎賢(KWAN Wing-yin, Wendy)	香港	香港政府資訊科技署
阿南康宏(Anan, Yasuhiro)	日本	微軟產品開發公司

姓 名	會 員 體	服 務 單 位
三橋洋一(Mihashi, Yoichi)	日本	大日本螢幕製造公司
山本知(Yamamoto, Satoshi)	日本	日立公司
川幡太一(Kawabata, Taichi)	日本	日本電訊公司
山本太郎(Yamamoto, Taro)	日本	Adobe系統公司
朴鍾寓(Park, Jong-woo)	南韓	高麗大學韓國文化研究所
Kim, Kyongsok	南韓	國立釜山大學
Ahn, Byoung-Ho	南韓	忠清大學
Bae, Jin-Seok	南韓	技術標準局
吳中越(Ngo, Trung Viet)	越南	越南資訊技術研究所
蘇重德(To, Trong Due)	越南	社會科學與人文中心漢喃研究所
畢夏普(Bishop, Tom)	美國	Wenlin Institute, Inc
小林劍(Lunde, Ken)	美國	Adobe 系統公司
Eric Muller	美國	Adobe 系統公司
Rick McGowan	美國	Unicode 協會
Magda Danish	美國	Unicode 協會
曲理察(Cook, Richard S)	美國	柏克萊加州大學語言學系
Ksar, Mike	美國	WG2主席
井作恒(Jenkins, John)	Unicode協會	Apple電腦公司

五、會議目的

為容納全球各種語言的字元和符號，ISO國際標準組織的會員國於1984年組成ISO/IEC JTC1/SC2/WG2（簡稱WG2），負責制定新的國際字元集編碼標準“Universal Multiple-Octet Coded Character Set”（簡稱UCS），編號則訂為ISO/IEC 10646（簡稱ISO 10646）；其中「中日韓認同表意文字區」即為收容亞洲各國所使用之漢字。因漢字集的規模龐大，且為多個國家（地區）共同使用，WG2為此再設「表意文字小組」（Ideograph Rapporteur Group,簡稱IRG），專責蒐集、比對各國（地區）之中文字集，再向WG2提出彙整後漢字集；我國自73年起即由台北市電腦公會代表參與該標準草案之編訂，嗣後由本局專案委託財團法人資訊工業策進會、中央研究院、財團法人中文數位化技術推廣基金會等辦理將我國CNS 11643「中文標準交換碼」資訊用字納入ISO 10646國際標準中，並代表出席相關會議。CNS 11643「中文標準交換碼」分為16字面，其中第1～7字面列有教育部所公布的常用、次常用、罕用及異體等字集，第8～11字面預留空白，而12～16為戶政等系統之使用者加字區，多年來經各相關單位，CNS 11643第1～7字面的中文字集，已全數編入ISO 10646，包括ISO 10646-1「廣用多八位元編碼字元集（UCS）－基本多語文字面」計23,989個字及ISO 10646-2「廣用多八位元編碼字元集（UCS）－輔助字面」計29,950個字。

本次會議之目的為審查CJK_D字集，討論急需編碼字集、專屬用字編碼區及表意文字認同規則，擬訂IRG下一階段工作方向，並確認IRG第30～32次會議舉辦地點與時間。

貳、會議過程

◆ 11月11日（星期日）：去程

◆ 11月12日（星期一）上午：

- 1、主辦單位至歡迎詞。
- 2、主席宣布開會並說明本次會議重要待議事項。
- 3、與會人員逐一自我介紹(roll call)。
- 4、確認議程，議程請詳見IRG N1340 “IRG#29 Meeting Agenda”。
- 5、主席指派大會秘書及決議起草小組（各會員體派一名代表）。
- 6、各會員體代表報告各自相關活動，我國的活動報告請詳附件IRG N1362 “Activity Report form TCA”。
- 7、首席編輯報告自IRG第28次會議決議的執行情形，以及WG2第51會議與IRG相關的決議項目。
- 8、古漢字工作組報告，請詳見IRG N1345 “Consolidated Old Hanzi of Sept meeting”。
- 9、日本提案報告。

◆ 11月12日（星期一）下午：

1、CJK工作組：

- （1）討論WG2第51次會議決議中有關CJK_C決議項目（詳見IRG N1352）。
- （2）討論急需編碼字集議題：先由各會員體代表報告急需編碼字數及理由，然後投票決定可接受成爲急需編碼字集的來源字集及字數。

日本Adobe：21個字

台灣：台客語常用字24個字

Unicode：17個字

中國：77個字

日本：120個字

合計259個字

- （3）逐字審查各會員體於兩輪審查中建議應予認同之漢字。

- 2、其他：依主席要求，緊急提供24個台客語新增字編碼所需資訊給首席編輯，以便轉交技術編輯製作匯集字表供審查。

◆ 11月13日（星期二）上午：

CJK工作組：繼續逐字審查各會員體於兩輪審查中建議應予認同之漢字（參見IRG N1353“CJK_D_V22-D_Unification”），有疑義者保留。

◆ 11月13日（星期二）下午：

1、CJK工作組：繼續逐字審查各會員體於兩輪審查中建議應予認同之漢字（參見IRG N1353“CJK_D_V22-D_Unification”），有疑義者保留。

2、其他：依技術編輯要求，緊急修訂24個台客語新增字編碼所需資訊。

◆ 11月13日（星期二）晚上：大會晚宴。

◆ 11月14日（星期三）上午：

CJK工作組：

（1）討論保留的認同字，我國就被保留部分的回應，詳見IRG N1385“TCA Response”，其中有如下4組被列為討論認同規則之案例。

萱 06985 TD-313D	萱 8413	暇 03665 TD-3B6B	暇 U+6687	澳 04501 TD-3D7E	漁 U+06F01	罇 07423 TE-3062	罇 87F3
-----------------------	-----------	-----------------------	-------------	-----------------------	--------------	-----------------------	-----------

（2）討論我國所提修訂CJK_C中14個字排序屬性，以及編輯過程中所漏失3個我國文字之回復要求，詳見IRG N1379“TCA_CJK_C V10”。

◆ 11月14日（星期三）下午：

1、CJK工作組：繼續討論保留的認同字。

2、大會：

（1）Adobe公司技術人員Ken Lunde介紹“Ideographic variation sequence”。

（2）Tom Bishop和Richard Cook介紹CDL（字元描述語言）。

（3）Mike Ksar說明WG2概況。

3、表意文字認同規則修訂小組：

John Jenkins代表UTC建議IRG制訂作業原則和程序，詳見IRG N1382“UTC Recommendations for IRG Principles and Procedures”。主席隨即指派IRG Principles & Procedures起草小組、審查組，並安排期程。

◆ **11月15日（星期四）上午：**

1、CJK工作組：

- （1）繼續討論保留的認同字。
- （2）討論CJK_B多欄字表、CJK-D字集及急需編碼字集的審查時程。
- （3）起草CJK_D工作報告，詳見IRG N1388 “Editorial Group Report”。

2、表意文字認同規則修訂小組：

配合不在現場的John Knightley，以視訊會議方式繼續討論表意文字認同規則修訂草案，詳見IRG N1350 “Consolidated Comments on Annex S Revision”。

◆ **11月15日（星期四）下午：**

1、CJK工作組：

- （1）繼續討論保留的認同字。
- （2）繼續起草並通過工作報告，詳見IRG N1388 “Editorial Group Report”。

2、表意文字認同規則修訂小組：

起草並通過工作報告，詳見IRG N1397 “Annex S Review Ad Hoc Group Report”。

3、大會：

- （1）起草大會決議。決議包括指派大會主席將我國所提修訂CJK_C中14個字之排序，以及要求回復編輯過程中漏失的3個我國文字等事轉提交WG2處理。
- （2）討論並通過大會決議，詳見IRG N1377 “IRG#29 Resolution”。

◆ **11月16日（星期五）全天：主辦單位安排參訪。**

參、會議議題及決議

一、CJK字集之整理

1、CJK_D字集整理工作的時程如下：

完成日期	工作項目
2007-12-27	IRG 首席編輯分送 CJK_D 第 3 版給各會員體。
2008-01-25	IRG 各會員體將 CJK_D 第 3 版審查結果提交 IRG 首席編輯。
2008-02-15	IRG 首席編輯分送 CJK_D 第 3 版給 IRG 技術編輯及各會員體。
2008-02-29	IRG 技術編輯分送 CJK_D 第 3 版給各會員體。
IRG 第 30 次會議	於大會中討論 CJK_D 第 3 版。

2、CJK_B多欄字碼表審查工作時程如下：

完成日期	工作項目
2008-03-14	各會員首席編輯提交 CJK_B 多欄字碼表審查結果給 IRG 首席編輯。
2008-03-28	IRG 首席編輯將前述審查結果彙整後分送 IRG 各會員體。
2008-04-04	各會員首席編輯提交 CJK_B 多欄字碼表審查結果給 IRG 首席編輯。
2008-04-11	IRG 首席編輯將審查結果送交 IRG 技術編輯。
2008-04-18	IRG 技術編輯將前述審查結果彙整後分送 IRG 各會員體。
2008-06-02	各會員首席編輯提交 CJK_B 多欄字碼表審查結果給 IRG 首席編輯。
IRG 第 30 次會議	於大會中討論 CJK_B 多欄字碼表。

3、急需編碼字集工作時程如下：

完成日期	工作項目
2008-02-15	會員首席編輯提交急需編碼字集審查結果及 IDS 分析給 IRG 首席編輯。
2008-02-29	IRG 首席編輯將前述審查結果彙整後分送 IRG 各會員體。
2008-03-07	IRG 各會員體確認急需編碼字集。
2008-03-14	IRG 首席編輯分送急需編碼字集第 1.1 版給 IRG 技術編輯及各會員體。
2008-03-21	IRG 技術編輯分送急需編碼字集第 1.1 版給各會員體。
IRG 第 30 次會議	於大會中討論急需編碼字集第 1.1 版。
IRG 第 30 次會議	於大會中討論 CJK_B 多欄字碼表。

二、古漢字之整理

古漢字集整理工作時程如下：

完成日期	工作項目
2008-03-28	完成《說文解字》第 180 至第 260 部首之甲骨文並提交李國英教授。
2008-04-10	彙整完成之甲骨文分送工作組成員。
2008-06-02	將工作組成果提交 IRG 主席。

三、IRG規章起草

IRG規章起草工作時程如下：

完成日期	工作項目
2007-12-21	John Jenkins 將規章第 1 版分送各會員體。
2008-1-31	各會員體回應審查意見。
2008-2-29	John Jenkins 將規章第 2 版分送各會員體。
2008-3-31	各會員體回應審查意見。
2008-4-11	John Jenkins 將規章第 3 版分送各會員體。
2008-5-11	各會員體回應審查意見。
2008-5-26	規章第 4 版完成。

四、表意文字認同規則修訂

表意文字認同規則修訂工作時程如下：

完成日期	工作項目
2007-12-31	提供表意文字認同及不認同之表格範例（TCA），俾供討論。
2008-01-31	完成認同規則修訂第 1 版，並分送各會員體。
2008-03-15	各會員體提交審查意見給 John Knightley。
2008-04-15	完成認同規則修訂第 2 版，並分送各會員體。
2008-05-15	各會員體提交審查意見給 John Knightley。
2008-06-02	完成認同規則修訂第 2 版，並於 IRG #30 討論。

三、未來議程

1、IRG第30次會議：由南韓主辦。

地點：釜山。

日期：2008年6月9日～13日。

2、IRG第31次會議：由中國主辦。

地點：重慶。

日期：2008年11月17日～21日。

3、IRG第32次會議：暫訂香港主辦。

地點：香港城市理工學院。

日期：2009年6月。

肆、我國應配合辦理之工作

一、CJK字集的整理：

- 1、審查IRG所分送CJK_D第3版字集，並依規定時程提交IRG首席編輯。
- 2、審查IRG所分送CJK_B多欄字碼表，並依規定時程提交IRG首席編輯。
- 3、確認急需編碼字集，並審查IRG所分送急需編碼字集第1.1版，並依規定時程提交IRG首席編輯。

二、古漢字的整理：

- 1、提交《說文解字》第180-260之甲骨文部首字形資料。
- 2、審查IRG所分送甲骨文部首字形資料，並依規定時程提交IRG首席編輯。

三、IRG規章之起草

- 1、審查IRG規章第1版，並依規定時程提交IRG首席編輯。
- 2、審查IRG規章第2版，並依規定時程提交IRG首席編輯。
- 3、審查IRG規章第3版，並依規定時程提交IRG首席編輯。

四、表意文字認同規則之修訂

- 1、製作表意文字認同及不認同之表格範例，並提供IRG首席編輯，俾供各會員體參考。
- 2、審查表意文字認同規則第1版，並依規定時程提交IRG首席編輯。
- 3、審查表意文字認同規則第2版，並依規定時程提交IRG首席編輯。

上述應配合辦理之工作已促請本局96年度「中文資訊標準之維護與推廣」委辦計畫之受託單位中文數位技術基金會配合辦理。

伍、心得與建議

- 一、多年來經各相關單位努力，CNS 11643第1~7字面的中文字集，已全數編入ISO 10646，包括ISO 10646-1「廣用多八位元編碼字元集（UCS）—基本多語文字面」計23,989個字及ISO 10646-2「廣用多八位元編碼字元集（UCS）—輔助字面」計29,950個字。後續IRG將再擴編CJK_D字集及各會員急需編碼之字集，並修訂表意文字認同規則及草擬IRG規章。有關上述工作需配合審查及提供資料，已請「中文資訊標準之維護與推廣」委辦案之受託單位財團法人中文數位化技術推廣基金會配合辦理。
- 二、本次會議中，當討論到某個我國中文字應否與另一個字認同時，日本代表認為兩個字的結構相同，爰主張應予認同，我國代表則持反對意見，理由為依「ISO/IEC 10646附件S所規定」，結構不同之異體字不應認同，該等字不該被解釋成「結構相同之異體字應認同」，後經CJK工作組充分討論後，決議依我國代表的主張不予認同。由此可見，出席如IRG等工作會議時，必須具備足夠的專業知識，以備當工作組討論有關中文字是否應予認同時，可適時提出理由及事證。
- 三、本次會議於美國加州舊金山灣區Adobe公司舉行，Adobe的視訊會議有18支小型麥克風從天花板垂下，12個喇叭嵌入天花板，加上兩個可以一起或分別投影的銀幕，讓會議室中每個人的發言都可以被清晰的收音，同時清晰的收聽遠端的發言，兩個銀幕則可同時顯示討論主題和參考資料。比照國內的會議室，多半將麥克風和喇叭組固定安裝於會議桌上，此會造成非正前方的人發言不便，爰若能採用Adobe的設計，將可解決此問題。
- 四、ISO 10646可說是少數我國能夠積極參與制訂與維護的國際標準之一，後續IRG工作組將再持續進行有關漢字集的擴編、中文字的屬性及其古漢字的整理等多項工作，爰我國應能長期參與該工作組活動，並積極貢獻，才能建立及維持我國在該領域的影響力。

陸、附件

- 1、會議照片集錦
- 2、IRG N1340：IRG#29 Meeting Agenda（本次會議議程）
- 3、IRG N1345：Consolidated Old Hanzi of Sept meeting（古漢字工作組報告）
- 4、IRG N1352：Resolutions of WG2 meeting 51
- 5、IRG N1362：Activity Report form TCA（我國活動報告）
- 6、IRG N1353：CJK_D_V22-D_Unification
- 7、IRG N1385：TCA Response
- 8、IRG N1379：TCA_CJK_C V10
- 9、IRG N1382：UTC Recommendations for IRG Principles and Procedures
- 10、IRG N1388：Editorial Group Report
- 11、IRG N1350：Consolidated Comments on Annex S Revision
- 12、IRG N1397：Annex S Review Ad Hoc Group Report
- 13、IRG N1377：IRG#29 Resolution

附件 1

會議照片集錦

圖1 會議地點：Adobe公司總部



圖2 會議進行場景



圖3 我國代表



圖4 視訊會議情形



圖5 會議室中懸垂的麥克風和嵌入天花板的喇叭



附件 2

IRG N1340 : IRG#29 Meeting Agenda

(本次會議議程)

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N [1340](#)

Date: 2007-07-11

Source:	Lu Qin, IRG Rapporteur
Title:	1 st Call for IRG#29 meeting to be held in San Jose, California, U.S.A, 2007-11-12~2007-11-16
Status :	Tentative
Actions required	Travel arrangement, review and feedback of agenda items
Distribution:	
Medium :	IRG Members and Ideographic Experts
	Electronic

Dear IRG Members and Experts,

According to the resolution IRG M28.1 approved by IRG#28 in Xian, Shaanxi, China, the 29th meeting of the Ideographic Rapporteur Group (IRG) will be held in San Jose, California, U.S.A. from 2007-11-12 to 2007-11-16.

Location:

Adobe, Albertus conference room, 345 Park Avenue, San Jose, CA, U.S.A.

Date/Time:

IRG #29 will start on Monday, 12 Nov. 2007 at 10am¹ and conclude on Friday, 16 Nov, 2007.

IRG members are requested to review the tentative agenda and send comment to IRG Rapporteur anytime before Monday 5 Nov, 2007.

Lu Qin

IRG Rapporteur

FAX: (852) 2774 0842, TEL: (852) 2766 7274, Mobile: (852) 9684 0623

Email: csluqin@comp.polyu.edu.hk

¹ All status reports should be submitted to the IRG Rapporteur by Hong Kong time 9:00am on 8 Nov., 2007 (the Thursday before the meeting commencement) to guarantee distribution and presentation for the morning session on the 1st day of the meeting.

Tentative Agenda

1. Opening speech by host
2. Administrative Items
 - 2.1. Roll call
 - 2.2. Approval of agenda items
 - 2.3. Assignment of meeting secretary and drafting group, other duties
 - Meeting secretary (to take note)
 - Drafting group leader and members
 - Name card Collector & Address Book Revise
 - Participant List,
 - eMailing List
3. Review of follow up actions:
 - Review of resolutions and action items defined by IRG#28.
 - Report of WG2 resolutions related to IRG work
4. Members activity reports(Please submit report by **Thursday 8 November, 2007** to guarantee distribution):
 - Standardization and implementation
 - Working report of member's Chief Editor
5. Discussion and work items in IRG #29
 - 5.1 The Status of CJK_C submission and feedback from WG2
 - 5.2 CJK_D work
 - 5.3 Old Hanzi Interest Group
6. IRG Work Planning and Future Meetings
7. Editorial/Ad hoc group Meetings
 - 7.1 CJK-C follow up actions
 - 7.2 CJK-B Source Visual Reference
 - 7.3 CJK-D work
 - 7.4 Old Hanzi
8. Other Business
9. Closing : Approve Resolutions

LOGISTICS
ISO/IEC JTC1/SC2/WG2/IRG Meeting # 29 San Jose,
CA – USA Nov 12-15, 2007

MEETING PLACE

Adobe

345 Park Avenue, San Jose, CA 95110, U.S.A.

[See Attached Map from Google Search\(AppendixA\).](#)

The meeting will take place in the Albertus conference room located next to the lobby of the Adobe building entrance at 345 Park Avenue

Free parking will be available. Let parking attendant and security officers know that you are a participant of the IRG meeting hosted by Eric Muller at Adobe.

CONTACT PERSON AT ADOBE

Eric Muller emuller@adobe.com Phone number: 408-536 4085 During the meeting only or in case of emergency, Eric can be contacted at 650-906-8135

DIRECTION TO MEETING PLACE FROM AIRPORT

Please see Page one of the attached [Adobe Customer Guide from Adobe\(Appendix B\)](#) for direction to the meeting place.

HOTELS AND RESTAURANTS

Please see Page 2 of the attached [Adobe Customer Guide from Adobe\(Appendix B\)](#) for hotel listings.

Delegates are kindly requested to make their own hotel reservation. They can ask for the Adobe corporate rate.

MEETING REGISTRATION FORM

Kindly complete the Meeting Registration Form as indicated and fax or email back to:

Magda Danish, magda@unicode.org or Fax # +1-650-693-3010.

See attached Sample Invitation Letter from Unicode([Appendix A](#) of IRGN1339)

Participant Information

Last Name	
First Name	
Organization	
Title	
Address	
City	
Postal Code	
Country	
Telephone	
Fax	
E-mail	

Accompanying Guests Information

	Last Name	First Name
1.		
2.		

Date: _____ Signature: _____

附件 3

IRG N1345: Consolidated Old Hanzi of Sept meeting

(古漢字工作組報告)

**Universal Multiple-Octet Coded Character Set
UCS**

ISO/IEC JTC1/SC2/WG2 IRG N 1345

Date: 2007-9-1

Title:	Consolidated Old Hanzi Submission (Radical 101 to 180 in Shouwen)
Source:	IRG N1303(TCA), IRGN 1307(China)
Status:	Input to IRG
Action:	
Distribution:	IRG Members and Ideographic Experts
Reference:	
No. of pages:	1
Medium:	Electronic

This submission document is according as the consolidation of two documents. Both of the documents are TCA IRG N1343 and China IRG N1344.

The method of integrating the data from these two documents is using some ordinal key fields to sort all the data in one format. First of all, using the filed of SW Radical Number to group all the relate records. If there are too many records relate to one SW Radical Number, and then will use the second field which is Corresp. Modern Character, to sort again. Finally, if after using the above two keys still couldn't organize all the records into a fine structure, then will use "ID" field to be the last way to order the rest of the records.

Documents submitted to the IRG #29 Meeting are listed below:

The electronic files are a PDF file, for Oracle Bone inscriptions for reviewing (IRGN1345-Appendix A1-Consolidated Old Hanzi Submission).

End of document

附件 4

IRG N1352 : Resolutions of WG2 meeting 51

**Universal Multiple-Octet Coded Character Set
(UCS)**

IRGN 1352

ISO/IEC JTC 1/SC 2 N3972R2

ISO/IEC JTC 1/SC 2/WG 2 N 3354R2

Date: 2007-10-27

Source: WG 2 meeting 51, Hangzhou, China, 2007-09-17/21
Title: Resolutions of WG 2 meeting 51
Action: For approval by SC 2 and for information to WG 2
Status: Adopted at meeting 51 of WG 2
Distribution: ISO/IEC JTC 1/SC 2 and WG 2

Experts from Canada, China, Finland, Ireland, Japan, Korea (Republic of), Poland, SEI - UC Berkeley (Liaison), Taipei Computer Association (Liaison), UK, Unicode Consortium (Liaison), and USA were present when the following resolutions were adopted (see attached attendance list).

Character count 99016 (till end of Amd. 2)

1102 additions in FDAM3

636 additions in FPDAM4

5719 additions in PDAM5

Total: 106473 (prior to meeting M51)

RESOLUTION M51.1 (Myanmar changes):

Unanimous

With reference to document N3277, WG2 accepts the rearranged code chart for Myanmar block including the following changes (in the new code positions):

- a. Change the glyphs for 107A MYANMAR LETTER SHAN NYA
- b. Change the name for 108C to MYANMAR SIGN SHAN COUNCIL TONE-3, and adjust the base line for the glyph
- c. Add the following 17 characters:

1077 MYANMAR LETTER SHAN GA	1093 MYANMAR SHAN DIGIT THREE
1079 MYANMAR LETTER SHAN ZA	1094 MYANMAR SHAN DIGIT FOUR
107B MYANMAR LETTER SHAN DA	1095 MYANMAR SHAN DIGIT FIVE
107F MYANMAR LETTER SHAN BA	1096 MYANMAR SHAN DIGIT SIX
108B MYANMAR SIGN SHAN COUNCIL TONE-2	1097 MYANMAR SHAN DIGIT SEVEN
1090 MYANMAR SHAN DIGIT ZERO	1098 MYANMAR SHAN DIGIT EIGHT
1091 MYANMAR SHAN DIGIT ONE	1099 MYANMAR SHAN DIGIT NINE
1092 MYANMAR SHAN DIGIT TWO	109E MYANMAR SYMBOL SHAN ONE
	109F MYANMAR SYMBOL SHAN EXCLAMATION
- d. Remove the tiny dot in the glyph for 1040 MYANMAR DIGIT ZERO

The rearranged code chart and names list are shown on page 7 of document N3277 as updated in this meeting.

count: +17 =653 in Amd.4

RESOLUTION M51.2 (Lanna script):

Unanimous

With reference to the ad hoc report on Lanna script in document N3346, and based on disposition of ballot comments on FPDAM4 at the meeting, WG2 resolves to move all the changes introduced for Lanna script (block 1A20--1AAF) from Amendment 4 to Amendment 5, incorporating the following changes:

- a. Ensure the correct glyphs and associated names are used to produce the charts correcting the previous production error
- b. Replace LANNA with TAI THAM in the script name, the block name and all the character names
- c. Add an annotation that the script is also known as LANNA (similar to TAI LUE)
- d. Insert a new character 1A5A TAI THAM SIGN LOLA HOY KONGNOT
- e. Move the characters in 1A5A to 1A5E down one position
- f. Move the character in 1A5F to 1A7E.

Further the Chinese national body and other experts are invited to discuss and resolve all outstanding concerns from document N3346 for consideration at the next WG2 meeting.

count: -127= 526 in Amd.4

count: +128= 5847 in Amd.5

RESOLUTION M51.3 (Disposition of FPDAM4 ballot comments):

Unanimous

WG2 accepts the disposition of ballot comments on FPDAM4 in document N3344 and instructs its editor to prepare the final text of Amendment 4 incorporating the dispositions. The following changes are noted in particular:

- a. Lanna script is moved out to Amendment 5 per resolution M51.2 above
- b. Improvement of the glyph for 1E9E
- c. Correction to the glyphs for 0333 and 0347
- d. Addition of annotations to MALAYALAM LETTER CHILLU RR and CHILLU L
- e. Changes in Myanmar Shan entries per resolution M51.1 above
- f. Rearrangement of characters in columns 2E2, 2E3 and 2E4 in the Supplemental Punctuation block per Irish ballot comment T.4 and accompanying modified chart on page 5 in document N3330
- g. Replacement of 'SWUNG DASH' with 'TILDE' in the names for 2E1B, 2E1E and 2E1F
- h. Corrections to several source reference errors for CJK
- i. Corrections to collection number allocations
- j. Removal of 2018--201F, 301D-301F from the BiDi mirrored list

count = 526 in Amd. 5

RESOLUTION M51.4 (Progression of Amendment 4):

Unanimous

WG2 resolves to include all the items accepted for inclusion in the standard noted in resolutions M51.1 and M51.2 into Amendment 4. WG2 instructs its project editor to forward the final text of Amendment 4 along with the disposition of comments document N3344 to the SC2 secretariat for an FDAM ballot. The final set of charts and names lists are in document N3348. The unchanged target starting date for FDAM4 is 2007-12.

*Count = net 526 additions in FDAM4
Total: 100118+526 = 100644 till end of FDAM4*

RESOLUTION M51.5 (Removal of Bamum script from Amd. 5):

Unanimous

With reference to document N3298, and supportive technical comments in PDAM5 ballot response document N3331, WG2 resolves to remove all changes introduced for Bamum script (block 1C80--1CCF) from Amendment 5, awaiting further input from the Bamum user community before further processing.

count: -88= 5759 in Amd.5

RESOLUTION M51.6 (Removal of Meitei Mayek script from Amd. 5):

Unanimous

With reference to document N3327, and based on discussion of PDAM 5 ballot comments on Meitei Dandas at the meeting, WG2 resolves to remove all changes introduced for Meitei Mayek script (block A6A0--A6FF) from Amendment 5, until the question about script-specific Dandas is resolved (see resolution M51.7 below).

count: -76= 5681 in Amd.5

RESOLUTION M51.7 (Criteria for script-specific Dandas):

Unanimous

WG2 invites the German, Indian, Irish, UK and US national bodies in consultation with other experts, to formulate 'definitive criteria for script-specific Dandas' for consideration at the next WG2 meeting for inclusion in the WG2 Principles and Procedures.

RESOLUTION M51.8 (Word Separator Middle Dot)

Unanimous

With reference to documents N3347 and Irish ballot comment T.3 on PDAM 5 in document N3331, WG2 accepts to encode

2E31 WORD SEPARATOR MIDDLE DOT

with its glyph as shown in document N3349.

count: +1 =5682 in Amd. 5

RESOLUTION M51.9 (Update to 11 KP source references)

Unanimous

WG2 accepts to update the KP1 source reference mapping for 11 code points per document N3285.

RESOLUTION M51.10 (Japanese ARIB ideographs):

China - Abstention

Canada, Finland, Ireland, Japan, Korea (Republic of), Poland, UK and USA - Acceptance

With reference to document N3318, WG2 accepts to encode the following:

- 2 CJK unified ideograph characters: 9FC4 (ARIB ID 47), 9FC5 (ARIB ID 95), and
- 4 CJK Compatibility ideograph characters: FA6B (ARIB ID 39), FA6C (ARIB ID 67), FA6D (ARIB ID 93), and FA6E (ARIB ID 105)

with their glyphs and compatibility mappings as shown in Table 7-11 in document N3318.

Further, WG2 requests IRG to review and provide feedback on the above at its November 2007 meeting.

count: +6 = 5688 in Amd.5

RESOLUTION M51.11 (Disposition of PDAM5 ballot comments):

Unanimous

WG2 accepts the disposition of ballot comments on PDAM5 in document N3345 and instructs its editor to prepare the final text of Amendment 5 incorporating the dispositions. The following changes are noted in particular:

- a. Addition of Tai Pham script (also known as Lanna) per resolution M51.2 above
- b. Removal of Bamum script per resolution M51.5 above
- c. Removal of Meitei Mayek script per resolution M51.6 above
- d. Addition of 2E31 per resolution M51.8 above
- e. Updates to KP source reference per resolution M51.9 above
- f. Addition of 6 ideographs per resolution M51.10 above
- g. Removal of 67 CJK Ext. C unified ideographs (see also document N3280 from IRG)
- h. Corrections to 26 CJK Ext. C ideographs per ballot comments from Japan
- i. Corrections to glyphs for 14 CJK Ext. C unified ideographs (see also documents N3279 and N3281)
- j. Corrections to source references for several CJK Ext. C unified ideographs
- k. Deletion of 1 Egyptian hieroglyph at 130FB
- l. Addition of 9 Egyptian hieroglyphs
- m. Rearrangement of code chart for Egyptian hieroglyphs reflecting the above changes and removing any empty code positions
- n. Addition of missing annotations for Korean Hangul additions
- o. Cleanup of the source references along with IRG-provided G_KX data and hex notation in RoK source references.
- p. Updates to the names list for Hangul Jamos to avoid confusion
- q. Corrections to collection number errors

count:-56 = 5632 in Amd. 5

Correction: Item g above should read:

- g. Removal of 63 CJK Ext. C unified ideographs and removal of T-source for 4 CJK Ext. C unified ideographs (per charts in document N3280 attachment A). Further ideograph at 2AC85 has only a single reference, the T-source TD-3C3C; if this source reference is removed this ideograph will have no source reference and hence has to be deleted. The net will be removal of 64 ideographs and removal of only the T-source references for 3 ideographs from CJK Ext. C.

Disposition of comments N3345 should also reflect this correction.

(The running counts for Amd. 5 and the total have also been updated.)

RESOLUTION M51.12 (Progression of Amendment 5):

Unanimous

WG2 instructs its project editor to forward the final text of Amendment 5 along with the disposition of comments document N3345 to the SC2 secretariat for an FPDAM ballot. The final set of charts and names lists are in document N3349. The unchanged starting dates for this work item are: FPDAM 2007-10-31, and FDAM 2008-07.

count: net 5632 additions in FPDAM5

Total: 100644+5632 = 106276 till end of FPDAM5

RESOLUTION M51.13 (Imperial Aramaic script):

Unanimous

WG2 accepts to encode in the standard 31 characters in code positions 10840 to 10855 and 10857 to 1085F, in a new block 10840 to 1085F named Imperial Aramaic, with their names and glyphs as shown in document N3339.

count: 31 = 31 in Amd.6

RESOLUTION M51.14 (Phoenician numbers):

Unanimous

With reference to document N3284, WG2 accepts to encode the following 2 characters:

1091A PHOENICIAN NUMBER TWO
1091B PHOENICIAN NUMBER THREE
with their glyphs as shown on page 5 in document N3284.

count: +2 = 33 in Amd.6

RESOLUTION M51.15 (Inscriptional Parthian):

Unanimous

WG2 accepts to encode in the standard 30 characters in code positions 10B40 to 10B55 and 10B58 to 10B5F, in a new block 10B40 to 10B5F named Inscriptional Parthian, with their names and glyphs as shown on page 18 in document N3286.

count: 30 = 63 in Amd.6

RESOLUTION M51.16 (Inscriptional Pahlavi):

Unanimous

WG2 accepts to encode in the standard 27 characters in code positions 10B60 to 10B72 and 10B78 to 10B7F, in a new block 10B60 to 10B7F named Inscriptional Pahlavi, with their names and glyphs as shown on page 19 in document N3286.

count: +27= 90 in Amd.6

RESOLUTION M51.17 (Common Indic Number Forms):

Unanimous

With reference to documents N3312, N3334 and N3316, WG2 accepts to encode 10 characters in code positions A830 to A839 in a new block A830 to A83F called Common Indic Number Forms with glyphs as shown on page 14 of document N3316.

count: +10 = 100 in Amd.6

RESOLUTION M51.18 (Bengali Ganda Currency mark):

Unanimous

With reference to document N3311, WG2 accepts to encode:
09FB BENGALI GANDA MARK
with its glyph as shown in document N3311.

count: +1 = 101 in Amd.6

RESOLUTION M51.19 (Tibetan Religious symbols):

Unanimous

With reference to document N3268, WG2 accepts to encode the following 4 characters:

0FD5 TIBETAN SYMBOL GYUNG DRUNG NANG -KHOR
0FD6 TIBETAN SYMBOL GYUNG DRUNG PHYI -KHOR
0FD7 TIBETAN SYMBOL GYUNG DRUNG NANG -KHOR BZHI MIG CAN
0FD8 TIBETAN SYMBOL GYUNG DRUNG PHYI -KHOR BZHI MIG CAN

with their glyphs as shown on page 1 in document N3268.

count: +4 = 105 in Amd.6

RESOLUTION M51.20 (Amendment 6 – subdivision and PDAM text):

USA - Abstention

Canada, China, Finland, Ireland, Japan, Korea (Republic of), Poland, and UK - Acceptance

WG2 instructs its editor / convener to prepare a project sub division proposal (see document N3358) and PDAM text based on resolutions M51.13 to M51.19 above, and forward them to the SC2 secretariat for ballot. The proposed start dates for the progression of this work item are: PDAM 2007-11, FPDAM 2008-06, and FDAM 2008-11.

*count: net 105 additions in PDAM6
Total: 106276+105=106381 till end of PDAM6*

RESOLUTION M51.21 (FCD of next edition):

Unanimous

With reference to documents N3274, N3275 and N3276, WG2 instructs its editor / convener to prepare:

- a. a subdivision proposal for the next edition of the standard (see document N3360)
- b. text for the next edition of the standard incorporating the texts of Amendments 1 through Amendment 5

and submit the above along with the updated rationale document N3362 to SC2 for registration and balloting, with the following schedule: FCD: 2008-03 and FDIS: 2008-11.

Total of 106276 in 2nd edition

RESOLUTION M51.22 (Nüshu script)

Unanimous

With reference to documents N3287, N3322, N3337 and N3340 on Nüshu script, WG2 invites national bodies and liaison organizations to review and provide feedback to China and invites the Chinese national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.23 (Samaritan alphabet)**Unanimous**

With reference to document N3291 on Samaritan alphabet, WG2 invites national bodies and liaison organizations to review and provide feedback to Ireland and invites the Irish national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.24 (Javanese script)**Unanimous**

With reference to documents N3292, N3319 and N3329 on Javanese script, WG2 invites national bodies and liaison organizations to review and provide feedback to Ireland and invites the Irish national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.25 (Psalter Pahlavi script)**Unanimous**

With reference to document N3286 containing a proposal on Psalter Pahlavi script, WG2 invites national bodies and liaison organizations to review and provide feedback on Psalter Pahlavi script to Ireland and invites the Irish national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.26 (Book Pahlavi script)**Unanimous**

With reference to document N3294 on Book Pahlavi script, WG2 invites national bodies and liaison organizations to review and provide feedback to Ireland and invites the Irish national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.27 (Old South Arabian script)**Unanimous**

With reference to documents N3296 and N3309 on Old South Arabian script, WG2 invites national bodies and liaison organizations to review and provide feedback to Ireland and invites the Irish national body to work with the US national body and submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.28 (Tangut script)**Unanimous**

With reference to documents N3297, N3307, N3333 and N3343 on Tangut script (also known as Xi Xia script), WG2 invites national bodies and liaison organizations to review and provide feedback to the US national body and invites the US national body to work with Chinese, Irish and UK national bodies and invited experts and submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.29 (Old Turkic script)**Unanimous**

With reference to documents N3299 and N3357 on Old Turkic script (also known as Orkhun script), WG2 invites national bodies and liaison organizations to review and provide feedback to SEI and invites SEI to work with the Chinese national body and submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.30 (Old Lisu script)**Unanimous**

With reference to documents N3317, N3323 and N3326 on Old Lisu script, WG2 invites national bodies and liaison organizations to review and provide feedback to China and invites the Chinese national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.31 (Simple Miao script)**Unanimous**

With reference to document N3335 on Simple Miao script, WG2 invites national bodies and liaison organizations to review and provide feedback to China and invites the Chinese national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.32 (3 Malayalam characters)**Unanimous**

With reference to document N3295 on 3 Malayalam archaic characters, WG2 invites national bodies and liaison organizations to review and provide feedback to Ireland and invites the Irish national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.33 (Japanese TV Symbols)

Unanimous

With reference to document N3341 on Japanese TV symbols, WG2 invites national bodies and liaison organizations to review and provide feedback to US national body and invites the US national body to submit a revised proposal incorporating the feedback for consideration at WG2 meeting 52.

RESOLUTION M51.34 (Concern on delay in FDAM 3 progression):

Unanimous

WG2 instructs its convener to communicate to ITTF via SC2 secretariat and SC2 chair its concern on the delay by ITTF in processing FDAM3 to ISO/IEC 10646: 2003 with a request for expediting its processing.

RESOLUTION M51.35 (IRG Principles and Procedures):

Unanimous

With reference to item 3 in document N3283, WG2 invites IRG to develop a set of principles and procedures to guide their work, with due considerations for the current WG2 Principles and Procedures.

RESOLUTION M51.36 (IRG Annex S Review):

Unanimous

With reference to item 4 in document N3283, WG2 endorses the IRG activity to review and feedback on Annex S of ISO/IEC 10646: 2003, without impacting already unified ideographs in the standard, taking into consideration the FCD ballot progression which starts in 2008-03.

RESOLUTION M51.37 (IRG Feedback on Security and IICORE):

Unanimous

With reference to item 7 in document N3283, WG2 thanks the IRG for the feedback in document N3282. WG2 requests SC2 to forward this document to national bodies and liaison members for information.

RESOLUTION M51.38 (IRG ideographs for Names):

Unanimous

With reference to item 8 in document N3283, WG2 endorses the IRG activity to investigate and report back to WG2 on the issues and recommendations on ideographs for names of persons, places and the like.

RESOLUTION M51.39 (IRG Urgent-Need ideographs):

Unanimous

With reference to item 10 in document N3283, WG2 requests the IRG to report back to WG2 with a more complete plan related to the identified 12000 'urgent need' ideographs, along with a prioritization of this work with respect to other existing IRG work items.

RESOLUTION M51.40 (Roadmap snapshot):

Unanimous

WG2 instructs its convener to post the updated snapshot of the roadmaps (in document N3306) to the WG2 web site.

RESOLUTION M51.41 (Future meetings):

Unanimous

WG 2 meetings:

Meeting 52 - 2008-04-21/25, Redmond, WA, USA; SC2 plenary is planned for the 25th of April

Meeting 53 - Fall 2008, seeking host; backup USA

Meeting 54 - Spring 2009, Republic of Korea (pending confirmation); backup USA

Meeting 55 - Fall 2009 (along with SC2 plenary); Tokushima, Japan (pending confirmation); backup USA

RESOLUTION M51.42 (Appreciation to DKUUG for web site support):

By Acclamation

WG 2 thanks DKUUG and its staff, in particular Mr. Kristen Nielsen, for its continued support of the web site for WG 2 document distribution and the e-mail server.

RESOLUTION M51.43 (Appreciation to Host):**By Acclamation**

WG 2 thanks the Chinese national body (CESI), in particular Mr. Chen Zhuang, Ms. Huang Shanshan, Ms. Wang Li and Mr. He Zheng-an, for hosting the meeting, for providing excellent meeting facilities, and their kind hospitality.

Character count 99016 (till end of Amd. 2)
Additions: 1102 in FDAM3; 526 in FDAM4; 5632 in FPDAM5; and 105 in PDAM6
Total count: 106381 (after meeting M51)

Attendance List

The following **42** attendees representing **9** national bodies, **3** liaison organizations, including **4** invited experts were present at different times during the meeting.

Name	Representing	Affiliation
Mike KSAR	Convener, USA	Independent
HE Zheng-an	host, China	Chinese Electronics Standardization Institute
HUANG Shanshan	host, China	Chinese Electronics Standardization Institute
WANG Li	host, China	Chinese Electronics Standardization Institute
Adrian CHEUK	Invited expert, China	SIL International
Erich FICKLE	invited expert, China	SIL Internaional
Daniel ZHANG	Invited expert, Sweden	Sony Ericsson Mobile Communication
David MORSE	Invited expert, Thailand	LISU Literature outreach
LU Qin	IRG Rapporteur	Hong Kong Polytechnic University
Tatsuo KOBAYASHI	SC2 Chair, Japan	Justsystems Corporation
Deborah ANDERSON	SEI, UC Berkeley - Liaison	Dept. of Linguistics, UC Berkeley
Shih-Shyeng TSENG	TCA - Liaison	Academia Sinica
Alain LABONTÉ	Canada; Editor 14651	Ministère des services gouvernementaux du Québec
V. S. (Uma)	Canada; Recording Secretary	IBM Canada Ltd.
UMAMAHESWARAN		
Al Wendan	China	Xishuangbanna Newspaper Company
Al Xiang	China	Xishuangbanna Minority language Commission
CHEN Zhuang	China	Chinese Electronics Standardization Institute
HE Lifeng	China	Yunnan Minority Language Commission
HUANG Yanjun	China	Independent
LI Xulian	China	Minority Language Office, State Ethnic Affairs Commission
Mai Maitiaili	China	Xinjiang Province Committee of Language
ABUDUREYIMU		
RONG Fengmei	China	Yunnan Minority Language Commission
TAN Yuting	China	Yunnan Minority Language Commission
WANG Xiaoming	China	Institute of Applied Linguistics, Ministry of Education
Wushour SILAMU	China	Xinjiang University
XIE Ming yao	China	Jiangyong, Hunan
XIONG Yuyou	China	Yunnan Minority Language Commission
YIN Jianmin	China	Huaguang Company
YIN Jianghong	China	Beijing Founder Electronics Co. Ltd.
YU Kanglong	China	Xishuangbanna Newspaper Company
ZHAO Liming	China	Núshu Institute, Tsinghua University
ZHAO Shihong	China	Lijiang Dongba Culture Research Institute
ZHU Cuifang	China	Tsinghua University
Eero VITIE	Finland	CSC - Finnish IT Centre for Science
Michael EVERSON	Ireland; Contributing Editor	Evertime
Masahiro SEKIGUCHI	Japan	Fujitsu Limited
Dae Hyuk AHN	Korea (Republic of)	Microsoft Korea
KIM Kyongsok	Korea (Republic of)	Pusan National University
Elżbieta BROMA-WRZESIEN	Poland	Telekomunikacja Polska S.A.
John KNIGHTLEY	UK	Independent
Michel SUIGNARD	USA; Project Editor	Microsoft Corporation
Peter CONSTABLE	USA; The Unicode Consortium - Liaison; Contributing Editor	Microsoft Corporation

附件 5

IRG N1362 : Activity Report form TCA

(我國活動報告)

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1362

Date: 2007-11-7

Doc. Type:	Member body contribution
Title:	Activity Report form TCA
Source:	TCA
Status:	Input to IRG
Action:	For Your Information
Distribution:	IRG Members and Ideographic Experts
No. of pages:	2
Medium:	Electronic

1. 24 Taiwanese & Hakka characters are urgent need for encoding:

Taiwan's MOE (Ministry of Education) had collected 2,518 Taiwanese characters and 2,772 Hakka characters. Most of those characters are currently used in the textbooks for homeland courses in primary schools and junior high schools. From them, MOE found 33 new characters outside UCS. TCA double checked those 33 characters and found that nine characters should be unified. TCA proposed to add these 24 characters as urgent need characters for encoding. Please refer to the document IRG N1305 and its attachment.

2. The progression of old Hanzi work:

The old Hanzi expert group held a working meeting in Harbin, China on September 13th – 15th, 2007. They reviewed the consolidated old Hanzi set in IRG N1345. In addition, they discussed the document IRG N1346. The report from the meeting please refer to the document IRG N1354.

3. The standard “Sorting attributes for Chinese data processing” is being revised:

Based on the project from BSMI (Bureau of Standards, Metrology and Inspection, the organization for standardization in Taiwan), a standard named “CNS 14754: Sorting attributes for Chinese data processing” is being revised by CMEX (Chinese Foundation for Digitization Technology). CNS 14754 specified five attributes of Chinese characters: radix, stroke count without radix, total stroke count, stroke sequence and pronunciation. The attribute radix of a Chinese character is given by means of Kanxi radices. The attribute pronunciation of a Chinese character is presented in the form of Liu's code. Those attributes are intended used to sort formatted Chinese data, such as names of people. Because CNS 11643-1 specified a new classification to the basic stroke for

Chinese characters, the attribute stroke sequences must be revised according the new classification.

4. Award to www.cns11643.gov.tw:

The Electronic Data Processing center of the Directorate-General of Budget, Accounting and Statistics got the Gold Penguin Award because its long term effort to build and maintain the official web site for promoting CNS 11643 and relative standards. The Gold Penguin Award because was created for encouraging the developing and promoting works on the open software.

附件 6

IRG N1353 : CJK_D_V22-D_Unification

CJK D V2.2 D Unification

D_V20	U1CODE	U2CODE	U3CODE	U4CODE	G Answer	J Answer
SOURCE	U1_SOU.	U2_SOU.	U3_SOU.	U4_SOU.	H Answer	T Answer
					V Answer	M Answer
𠄎 00005 G_BK100002	𠄎 6C11				G: not unified,different component H: Not unified with 6C11	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠄎 00082 G_BK100010	𠄎 51E1	𠄎 5342	𠄎 2007D		G: not unified,different component H: Not unified with 51E1, 5342 or 2007D	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠄎 00195 G_BK100023	𠄎 9AD8				G: not unified,different component H: Not unified with 9AD8	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠄎 00303 G_BK100053	𠄎 50C0				G: not unified,different component H: Not unified with 50C0	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/ in accordance with ANNEX S
𠄎 00356 G_BK100068	𠄎 511E				G: delete it,maybe wrong glyph,G source only H: Not unified with 511E	J: not unifiable because of different number of component (see S.1.4.1). (*1)
𠄎 00505 G_BK100090	𠄎 D-00508 G_BK100091				G: not unified,different component H: Not unified with D-00508	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
凍 00551 G_BK100097	凍 5032	凍 202D6			G: not unified,obviously different glyph H: Unifiable with 51CD, provided that they have the same meaning and pronunciation; but not unified with 5032 or 202D6	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
凍 00551 G_BK100097	凍 51CD					
𠄎 00860 G_BK100165	𠄎 2A817				G: delete it. May be same character. G source only. H: Not unified with 2A817	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)

D_V20 U1CODE U2CODE U3CODE U4CODE
 SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
 H Answer
 V Answer




















J Answer
 T Answer
 M Answer

馬 00863 G_BK100166	厖 D-02734 G_BK100468	廌 5ECC			G: delete it. May be wrong character. G source only. H: Not unified with D-02734 but may be unified with 5ECC provided that they have the same meaning and pronunciation	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
収 00903 G_BK100173	収 53CE	収 20B1E			G: not unified,different component H: Not unified with 53CE or 20B1E	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
埝 01473 G_BK100260	埝 21292				G: unified H: Unifiable with 21292, provided that they have the same meaning and pronunciation	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
賓 02168 G_BK100355	賓 8CD3	賓 FA64			G: delete it. Wrongly simplified form. G source only. H: Not unified with 8CD3 or FA64	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3) T: unifiable/duplicate
兕 01010 G_BK100367	兕 D-02323 G_BK100366	兕 D-02323 G_BK100366			G: not unified,different component H: Not unified with D-02323 or 5492	J: unifiable/duplicate to u-5492. If any objection show more evidence. (*5)
兕 01010 G_BK100367	兕 D-02323 G_BK100366	兕 5492				
幌 02685 G_BK100455	幌 385B				G: not unified,different component H: Not unified with 385B	J: unifiable/duplicate. If any objection show more evidence. (*5)
往 02879 G_BK100502	往 U+224C2				G: not unified,different component	
擲 03400 G_BK100595	擲 640A				G: not unified. Different structure and components. H: Not unified with 640A	J: not unifiable because of different number of component (see S.1.4.1). (*1)
攬 03406 G_BK100597	攬 D-03405 UTC00790				G: unified H: Same as D-03405, unifiable	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unified/duplicate
導 03705 G_BK100650	導 26958				G: not unified,different component H: Not unified with 26958	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)

D_V20 U1CODE U2CODE U3CODE U4CODE
 SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
 H Answer
 V Answer

J Answer
 T Answer
 M Answer

 03706 G_BK100651	 65F1				G: not unified,different component H: Unifiable with 65F1, provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S 吞 & 吞; 刊刊
 03717 G_BK100656	 28FC0				G: not unified,different component H: Not unified with 28FC0	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
 04140 G_BK100736	 9B30				G: not unified. One more stroke. This is from KX, it is suggested to include all KX characters in UCS. H: Not unified with 9B30	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
 04238 G_BK100746	 6BBB	 23A8A			G: not unified. different component. H: Not unified with 6BBB or 23A8A	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
 04317 G_BK100774	 2AD86				G: not unified,different structure H: Not unified with 2AD86	J: not unifiable because of different relative positions of components (see S.1.4.2). (*2)
 04940 G_BK100878	 821D				G: not unified. Different component. H: Not unified with 821D	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
 05145 G_BK100942	 2AF5B				G: unified H: Not unified with 2AF5B	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
 05293 G_BK100971	 24D8F				G: not unified,different component H: Not unified with 24D8F	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
 05298 G_BK100972	 7603				G: not unified,different component. This is from KX, it is suggested to include all KX characters in UCS. H: Unificable with 7603 provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/ in accordance with ANNEX S


05530
 G_BK101011


77C7

G: unified
H: Unificable with 77C7 provided that they have the same meaning

J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)

D_V20 U1CODE U2CODE U3CODE U4CODE
 SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
 H Answer
 V Answer

J Answer
 T Answer
 M Answer

					and pronunciation	
砣 05608 G_BK101025	砣 D-05627 G_CYY01038				G: not unified,different component H: Unificable with D-05627 provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/ in accordance with ANNEX S
葱 06929 G_BK101368	葱 44E4				G: not unified,different component H: Unificable with 44E4 provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
赟 07003 G_BK101392	赟 D-07008 G_BK101394				G: not unified,different component H: Not unified with D-07008	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/duplicate
斂 07113 G_BK101424	斂 8539				G: not unified. Different component. H: Not unified with 8539	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
谿 07797 G_BK101656	谿 8C11				G: not unified,different component H: Unificable with 8C11 provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/ in accordance with ANNEX S
软 08203 G_BK101804	软 8F6A				G: not unified,different component H: Unificable with 8F6A provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
邨 08435 G_BK101923	邨 2B499				G: unified H: Unificable with 2B499 provided that they have the same meaning and pronunciation	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
钱 08775 G_BK102050	钱 94B1				G: unified H: Unificable with 94B1 provided that they have the same meaning and pronunciation	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
驮 09599 G_BK102451	驮 9A6E				G: unified H: Unificable with 9A6E provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S

D_V20 U1CODE U2CODE U3CODE U4CODE
SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
H Answer
V Answer

J Answer
T Answer
M Answer

點 10390 G_BK102945	點 D-10389 UTC00776				G: unified H: same as D-10389, unifiable	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/duplicate
龜 10524 G_BK103010	龜 4E80				G: not unified,different component H: Not unified with 4E80	J: unifiable/duplicate. If any objection show more evidence. (*5)
坳 01402 G_CH100248	坑 5748				G: not unified,different component H: Not unified with 5748	J: not unifiable because of different number of component (see S.1.4.1). (*1)
惹 03011 G_CH100531	惹 6178				G: not unified,different component H: Not unified with 6178	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
茱 03817 G_CH100679	茱 3B49	茱 2F8DE			G: unified H: same as 2F8DE, unifiable; and also unifiable with 3B49 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/duplicate
桀 03835 G_CH100686	桀 67FD	桀 23415			G: not unified,different component H: Not unified with 67FD or 23415	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
帶 03932 G_CH100709	帶 3BC2				G: not unified,different component H: Not unified with 3BC2	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
樓 04014 G_CH100725	樓 6A13				G: not unified. Different component. H: Not unified with 6A13	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3) T: unifiable/ in accordance with ANNEX S
滌 04496 G_CH100803	滌 D-04497 TD-3E68				G: not unified.	
笛 05218 G_CH100955	笛 753E	笛 2F936			G: Same as 2F936, unifiable; but not unified with 753E H: Same as 2F936, unifiable; but not unified with 753E	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
擧 03386 G_CH101317	擧 D-03332 TD-5C4B	U+64E7			G: unified	

D_V20 U1CODE U2CODE U3CODE U4CODE
SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
H Answer
V Answer

J Answer
T Answer
M Answer

葡 06928 G_CH101367	葡 24C08				G: unified H: unifiable with 24C08, provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)
谕 07746 G_CH101620	谕 D-07747 G_GH101621				G: not unified,different component H: Not unified with D-07747	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𪛗 10137 G_CH102744	𪛗 D-10138 G_GH102745				G: unified H: Not unified with D-10138	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
鸚 10176 G_CH102776	鸚 D-10177 G_HC302777				G: not unified,different component H: Not unified with D-10177	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𪛗 10362 G_CH102932	𪛗 2B757				G: not unified,different component H: Not unified with 2B757	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𪛗 01435 G_CH200254	𪛗 5763				G: not unified,need more research H: unifiable with 5763, provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
𪛗 03282 G_CH200574	𪛗 6455				G: not unified,different component H: Not unified with 6455	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
𪛗 03810 G_CH200676	𪛗 6748				G: unified H: Unifiable with 6748, provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)
𪛗 05035 G_CH200910	𪛗 D-10163 G_gjz00306				G: unified H: Unifiable with D-10163, provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/ in accordance with ANNEX S
𪛗 05602 G_CH201023	𪛗 7841	𪛗 2544B			G: not unified,different component H: Not unified with 7841 or 2544B	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)

D_V20 U1CODE U2CODE U3CODE U4CODE
SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
H Answer
V Answer

J Answer
T Answer
M Answer

纣 06356 G_CH201145	纣 D-06361 G_GH101150				G: not unified,different component H: Not unified with D-06361 but may be unified with 7EA8 if they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/ in accordance with ANNEX S
纣 06356 G_CH201145	纣 7EA8					
蓼 06992 G_CH201390	蓼 26C9E				G: not unified,different component H: Unifiable with 26C9E	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
蓼 07517 G_CH201527	蓼 2B33F				G: not unified,different component H: Unifiable with 2B33F	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
译 07731 G_CH201613	译 8BD1				G: not unified,different component H: Not unified with 8BD1	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
蹀 08090 G_CH201776	蹀 8E5B				G: not unified,different component H: Not unified with 8E5B	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
轻 08202 G_CH201803	轻 D-08205 G_BK101808				G: not unified,different component H: unifiable with D-08205 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
钊 08708 G_CH201994	钊 2B523				G: not unified,different component H: Not unified with 2B523	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
钊 08718 G_CH202003	钊 949E				G: not unified,different component H: Not unified with 949E	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
蛸 07326 G_CH301469	蛸 D-07328 G_BK101471				H: Unifiable with D-07328, provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/duplicate

D_V20 U1CODE U2CODE U3CODE U4CODE
 SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
 H Answer
 V Answer

J Answer
 T Answer
 M Answer

羸 08224 G_CH301826	羸 D-08230 G_zfy00526				G: not unified,different component H: Not unified with D-08230	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3) T: unifiable/ in accordance with ANNEX S
钛 08713 G_CH302000	钛 949B				G: not unified,different component H: Unifiable with 949B, provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/ in accordance with ANNEX S
鸚 10166 G_CH302768	鸚 D-10167 G_CH102769				G: not unified,different component H: Not unified with D-10167	J: unifiable/duplicate. If any objection show more evidence. (*5)
𪔐 02516 G_CH400409	𪔐 5D7D				G: not unified,different component H: Not unified with 5D7D	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
钲 08711 G_CH401998	钲 948D				G: not unified,different component H: Not unified with 948D	J: unifiable/duplicate. If any objection show more evidence. (*5)
钼 08766 G_CH402042	钼 94BF				G: not unified,different component H: Not unified with 948F	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𪔑 09261 G_CH402292	𪔑 D-09264 UTC00327				G: unified H: Same as D-09264, unifiable	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/duplicate
懷 03079 G_CY100539	懷 61E5				G: not unified. Different component. H: Not unified with 61E5	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
芑 06862 G_CY101336	芑 8291	芑 26AFE	芑 26B0A	芑 2F98F	G: unifiable with either 8291 or 2F98F, provided that they have the same meaning and pronunciation; but not unified with 26AFE or 26B0A H: unifiable with either 8291 or 2F98F, provided that they have the same meaning and pronunciation; but not unified with 26AFE or 26B0A	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/duplicate
元 00487 G_CYY00001	元 4E93	元 4EA3	元		G: not unified,obviously different glyph	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)

D_V20 U1CODE U2CODE U3CODE U4CODE
 SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
 H Answer
 V Answer

J Answer
 T Answer
 M Answer

			5143		H: Not unified with 4E93, 4EA3, 5143, 5927, 592A or 72AC	
兀 00487 G_CYY00001	大 5927	太 592A	犬 72AC			
谷 00958 G_CYY00004	谷 593B				G: not unified,different component H: Not unified with 593B	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
火 00081 G_CYY00023	火 706B	灬 706C			G: not unified,different component H: Not unified with 706B or 706C	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
了 00083 G_CYY00024	了 2193C				G: not unified,different component. Place name used character, no way to change it. H: unifiable with 2193C, provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
吞 00111 G_CYY00037	吞 2A952				G: unified H: unifiable with 2A952, provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
𠂔 00127 G_CYY00040	𠂔 D-01373 G_CYY00194				G: not unified,different component H: Not unified with D-01373	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠂔 00137 G_CYY00050	𠂔 D-05206 G_CYY00948	𠂔 4E6A			G: not unified,different component H: Not unified with D-05206 or 4E6A	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
水 00167 G_CYY00065	水 6C34	氷 6C3A			G: not unified,different component H: Not unified with 6C34 or 6C3A (*1)	J: not unifiable because of different number of component (see S.1.4.1).
立 00170 G_CYY00069	大 5927	太 592A	𠂔 215D3		G: not unified,obviously different glyph H: Not unified with 5927, 592A or 215D3	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)

非
00070
G_CYY00070

非
975E

G: not unified,different component

J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)

D_V20 U1CODE U2CODE U3CODE U4CODE
SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
H Answer
V Answer

J Answer
T Answer
M Answer

					H: Not unified with 975E	
仇 00218 G_CYY00075	仇 201A7				G: unified H: unifiable with 201A7 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)
仇 00224 G_CYY00078	仇 4EC7				G: not unified,different component H: unifiable with 4EC7 provided that they have the same meaning and pronunciation	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3) T: unifiable/ in accordance with ANNEX S
𠂇 00961 G_CYY00156	𠂇 20BB0				G: not unified,different component H: Not unified with 20BB0	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠂇 01033 G_CYY00174	𠂇 5544				G: unified H: unifiable with 5544 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5) T: unifiable/ in accordance with ANNEX S
𠂇 01372 G_CYY00193	𠂇 4E2E	𠂇 5720	𠂇 200D6	𠂇 2123E	G: not unified,different structure&component H: Not unified with 4E2E, 5720, 200D6, 2123E, 5721 or 2123D	J: not unifiable because of different number of component (see S.1.4.1). (*1)
𠂇 01372 G_CYY00193	𠂇 4E2E	𠂇 5721	𠂇 2123D	𠂇 2123E		
𠂇 01375 G_CYY00196	𠂇 34DA				G: not unified,different component H: Not unified with 34DA	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠂇 01377 G_CYY00197	𠂇 D-02615 G_WZ10044 1	𠂇 5DE9			G: not unified,different component H: Not unified with D-02615 or 5DE9	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠂇 01378 G_CYY00198	𠂇 21255				G: not unified,need more research H: Unifiable with 21255 provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4) T: unifiable/ in accordance with ANNEX S

𠂇 𠂇
01400 D-01391
G_CYY00212,V04-463D

G: not unified,different component
H: Not unified with D-01391
V: Un-unifiable due to different

J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)

附件 7

IRG N1385 : TCA Response

D_V20 U1CODE U2CODE U3CODE U4CODE
 SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
 H Answer
 V Answer

J Answer
 T Answer
 M Answer

照 04720 TD-2C24	𠄎 7138				H: Not unified with 7138	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠄎 04740 TD-2C2E	𠄎 2424B				H: unifiable with 2424B provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)
𠄎 04721 TD-2C35	𠄎 D-04723 G_zjw01249				H: Not unified with D-04723	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
焰 04741 TD-2C41	焰 7130				H: Not unified with 7130	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
𠄎 04722 TD-2C4B	𠄎 713D				H: Not unified with 713D	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠄎 05075 TD-2D39	𠄎 D-05096 TD-6037				H: Not unified with D-05096	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠄎 05625 TD-2E36	𠄎 787A				G: not unified,different component. Need TCA confirmation. H: unifiable with 787A provided that they have the same meaning and pronunciation	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
𠄎 05995 TD-2F44	𠄎 7B3F				H: Not unified with 7B3F	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
𠄎 06651 TD-304A	𠄎 26564	𠄎 2656F			H: Not unified with 26564 or 2656F	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
萱 06985 TD-313D	萱 8413				H: Not unified with 8413	J: unifiable/duplicate. If any objection show more evidence. (*5)

Not unified, because they have different component in the phonetic part.

be unified to 9EC4.

D_V20
SOURCE

U1CODE U2CODE U3CODE U4CODE
U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
H Answer
V Answer

J Answer
T Answer
M Answer

<p>黃 06957 TD-315B</p>	<p>黃 9EC3</p>	<p>黃 9EC4</p>			<p>H: Not unified with 9EC3 or 9EC4</p>	<p>J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)</p>
<p>蛸 07342 TD-3226</p>	<p>蛸 86D2</p>				<p>H: Not unified with 86D2</p>	<p>J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)</p>
<p>誣 07626 TD-3263</p>	<p>誣 46BE</p>	<p>誣 27976</p>			<p>H: unifiable with 27976 provided that they have the same meaning and pronunciation; but not unified with 46BE</p>	<p>J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)</p>
<p>勳 09786 TD-3467</p>	<p>勳 U+29d53</p>					
<p>博 00780 TD-3647</p>	<p>博 U+535a</p>				<p>G: not unified.</p>	
<p>嗔 01140 TD-372B</p>	<p>嗔 D-01144 TD-3669</p>				<p>H: unifiable with D-01144 provided that they have the same meaning and pronunciation</p>	<p>J: unifiable/duplicate. If any objection show more evidence. (*5)</p>
<p>嗽 01148 TD-3737</p>	<p>嗽 20E8B</p>				<p>H: Not unified with 20E8B</p>	<p>J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)</p>
<p>埭 01644 TD-376E</p>	<p>埭 21421</p>				<p>H: Not unified with 21421</p>	<p>J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)</p>
<p>嬪 01996 TD-3843</p>	<p>嬪 5AC3</p>				<p>H: Unifiable with 5AC3 provided that they have the same meaning and pronunciation</p>	<p>J: unifiable/duplicate. If any objection show more evidence. (*5)</p>
<p>尋 02279 TD-393B</p>	<p>尋 21B36</p>				<p>H: Not unified with 21B36</p>	<p>J: unifiable/duplicate. If any objection show more evidence. (*5)</p>
<p>彙 02868 TD-3A36</p>	<p>彙 5F59</p>	<p>彙 5F5A</p>	<p>彙 22465</p>		<p>H: unifiable with 5F5A provided that they have the same meaning and pronunciation; but not unified with 5F59 or 22465</p>	<p>J: unifiable/duplicate. If any objection show more evidence. (*5)</p>

be unified.

be unified to 5F5A.

D_V20 U1CODE U2CODE U3CODE U4CODE
 SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
 H Answer
 V Answer

J Answer
 T Answer
 M Answer

愼 03027 TD-3A56	愼 613C	愼 614E	愼 2F8A8	G: not unified. This is from KX, it is suggested to include all KX characters in UCS. H: unifiable with 613C, 614E and 2F8A8 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)	
upon discussion.						
暇 03665 TD-3B6B	暇 U+6687				G: unified.	
棗 03925 TD-3C45	棗 D-03984 TD-4D2D	U+68D7				
not be unified, because their upper components are different.						
椅 03959 TD-3D2F, V04-	椅 6905				H: Not unified with 6905 V: Un-unifiable due to different reading and systematically used of the right component in Vietnamese documents	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
澳 04501 TD-3D7E	漁 U+06F01				G: unified. Need T confirmation.	
not be unified, because their lower components are different.						
漾 04508 TD-3E6F	漾 6E94				H: Not unified with 6E94	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
燻 04756 TD-3E7B	燻 243B6				H: unifiable with 243B6 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)
be unified.						
爲 04855 TD-3F62	爲 U+7232					
臻 05080 TD-3F7E	臻 2AF30				H: unifiable with 2AF30 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)
椿 05784 TD-4177	椿 411D				H: Not unified with 411D	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
積 05866 TD-4240	積 7A39				H: Unifiable with 7A39	J: unifiable/duplicate. If any objection show more evidence. (*5)

D_V20 U1CODE U2CODE U3CODE U4CODE
 SOURCE U1_SOU. U2_SOU. U3_SOU. U4_SOU.

G Answer
 H Answer
 V Answer

J Answer
 T Answer
 M Answer

貝 04812 TE-2D7D	𧀮 24444				H: Not unified with 24444	J: not unifiable because of different structure of a corresponding component (see S.1.4.3). (*3)
not be unified, because the case of unifying their right components cannot be found.						
𧀮 07423 TE-3062	𧀮 87F3				H: Not unified with 87F3	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
be unified.						
𧀮 08668 TE-322A	𧀮 9424				H: Not unified with 9424	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
𧀮 08961 TE-3234	𧀮 49B0				H: Unifiable with 49B0 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)
𧀮 09247 TE-3258	𧀮 97DE				H: Not unified with 97DE	J: Not sure. Need proposer to submit more evidences to confirm its glyph shape. (*4)
𧀮 09820 TE-3327	𧀮 U+9be4					
𧀮 09812 TE-332E	𧀮 U+9bb8					
𧀮 10083 TE-3337	𧀮 9D75				H: Unifiable with 9D75 provided that they have the same meaning and pronunciation	J: not unifiable because of different relative positions of components (see S.1.4.2). (*2)
擻 03411 TE-346C	擻 2ABD4				H: Not unified with 2ABD4	J: not unifiable because of different relative positions of components (see S.1.4.2). (*2)
灑 04625 TE-355F	灑 D-04624 UTC00440				H: unifiable with D-04624 provided that they have the same meaning and pronunciation	J: unifiable/duplicate. If any objection show more evidence. (*5)
礫 05712 TE-364B	礫 2B01D				H: Not unified with 2B01D	J: not unifiable because of different number of component (see S.1.4.1). (*1)

附件 8

IRG N1379 : TCA_CJK_C V10

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1379

Date: 2007-11-14

Doc. Type:	Member body contribution
Title:	Comments on CJK_C V10
Source:	TCA
Status:	Input to IRG
Action:	To be discussed in the IRG#29
Distribution:	IRG Members and Ideographic Experts
No. of pages:	2
Medium:	Electronic

1. Incorrect ordering:

2B52C-2B539: totally 14 ideographs with simplified radical ‘door’ were encoded in an incorrect ordering. See attachment in below.

2. Missing characters:

TD-4421, TD-3C3C and TD-4D2B have been encoded in CJK_C V90 (2AC83, 2AC85 and 2AC87). The error report require to change their radical from ‘moon’ to ‘meat’.

Remark: We suggest to provide a list for all changes between CJK_CV90 and CJK_CV91, and CJK_CV91 and CJK_CV10 as a formal IRG document.

2B520	閔 TD-473D	2B530	13	鞫 V4-565D
2B521	閔 TD-902A	2B531	14	𨮒 V4-565Z
2B522	閔 閔 G_2AW01763 TD-902F	2B532	1	𨮒 V4-4024
2B523	𨮒 K5-0036	2B533	2	𨮒 V4-5649
2B524	𨮒 V4-5644	2B534	5	
2B525	閔 TD-7171	2B535	6	
2B526	𨮒 V4-5648	2B536	7	
2B527	𨮒 G_HC102194	2B537	9	
2B528	閔 TE-2069	2B538	10	
2B529	閔 TE-3040	2B539	11	
2B52A	閔 K-65675	2B53A		𨮒 G_2AW00649
2B52B	閔 TE-407A	2B53B		𨮒 TC-2061
2B52C	3 𨮒 V4-564A	2B53C		𨮒 TC-3842
2B52D	4 𨮒 V4-564B	2B53D		𨮒 G_2AW00650
2B52E	8 𨮒 V4-564C	2B53E		𨮒 𨮒 G_2AW00651 TC-4274
2B52F	12 𨮒 V4-564E	2B53F		𨮒 TC-4277

附件 9

**IRG N1382 : UTC Recommendations for
IRG Principles and Procedures**

UTC Recommendations for IRG Principles and Procedures

John H. Jenkins

14 November 2007

I've been giving some thought to IRG principles and procedures and have a couple of recommendations I'd like to make.

There are two specific areas I'd like to focus on: the IRG's *collective memory*, and the *amount of data available* on individual characters.

One of the features of the original work done by the IRG's predecessor, the CJK/JRG, was that it was reproducible. That is, given the same sources and the same principles, another group of people would (in theory) be able to do the same work and reach pretty much the same outcome. This was largely because the IRG's work was based on widely available standards and dictionaries.

As the sources for the IRG's work become more obscure, this aspect of our work is being lost. Not only is it unlikely that a second group of people could start out with the same data and end up with the same results, but it's very unlikely that even the IRG itself would end up with the same results. This is because we don't track the decisions we make—particularly the editorial decisions—in any readily available form.

This also means that the editorial committee will have to periodically revisit unification issues it has already decided. There are cases where we found component shapes which are sometimes unified and sometimes not. We can generally tell what decision was made in the past, but we can't tell why it was made.

I feel that this problem can best and most easily be resolved by the editorial committee making its minutes from past meetings available at least to IRG members. We would probably not want these to be documents available on the IRG's public web site, but they should nonetheless be made available for editorial committee members.

If editorial committee members had access to the minutes from past meetings, we could more easily determine what precedents to apply and save time currently spent recreating decisions that were made in the past.

The other issue has to do with the amount of data that IRG members are making available regarding the characters in their collections. Certainly this is true for the UTC characters. For most of the characters we have submitted to the IRG, we have specific source references as well as pronunciation data and definitions. This data is

currently available only internally to the UTC itself (and not even always there).

The IRG is already expecting members to provide more information about submitted characters. I think tendency is a good one and should be encouraged. What would be most useful would be for this data to be available to editorial committee members in a uniform way, or at least in documents which are maintained in a database or on a Web site such as the IRG's public site.

My point here is not so much that the IRG needs to require more data from member bodies, but that it needs to make sure there is a known, systematic way of retrieving that data in its current form after it has been received. This could be done as simply as having a special page on the IRG Web site where the initial proposals and the evidence documents are kept accessible. These documents could be kept updated and current by the various IRG members.

It would also be helpful if evidence documents were divided into documents with text-only data (pronunciation and definition data) and documents with scans of sources. This would make it easier to look up data in the middle of a meeting, because we wouldn't have to wait to search for text inside a PDF when it's only text we want to find.

附件 10

IRG N1388 : Editorial Group Repor

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2/IRG N1388

Date: 2007-11-15

Source:	CJK editorial group
Title:	CJK editorial group report of IRG#29, San Jose
Status :	
Actions required	To be approved by IRG plenary
Distribution:	IRG#29
Medium :	Electronic
Appendix:	2

Editors from China, Japan, R.O.Korea, Vietnam, TCA, Hong Kong SAR and US/Unicode Consortium met at IRG#29, 2007-11, San Jose, US. The editors discussed the following issues.

1. Urgently needed characters

According to IRG resolution M28.11, China, Japan, R.O.Korea, Vietnam, TCA and US/Unicode Consortium submit their urgently needed characters to IRG#29.

- China: IRGN1366
- Japan: IRGN1367
- R.O.Korea: IRGN1369
- Vietnam: IRGN1370
- TCA: IRGN1305
- US/Unicode Consortium: IRGN1364

The editors reviewed all these documents and voted for the scope of urgently needed characters. The voting results limited the scope of urgently needed characters to modern use and stable sets of characters. The voting results rejected IRGN1369 (R.O.Korea) and ancient books-use characters and people's names-use characters in IRGN1366 (China). The editors would like to appreciate Vietnam for her withdrawing IRGN1370 to reduce the size of urgently needed characters. The editors welcome China, R.O.Korea and Vietnam to re submit those characters at suitable time in future.

Current urgently needed characters:

- China: 77
- TCA: 24

- Japan: 120
- US/Unicode Consortium: 37

The editors agreed to process Urgently Needed Characters Set and CJK_D in parallel based on their individually agreed schedules.

A glyph list with characters' attributes (IRGN1380) was generated by Ms. Wang Xiaoming and all IDSs were checked (IRGN1380 Appendix) by Mr. Taichi Kawabata during the meeting. The editors set the following schedule for review of urgently needed characters according to the quick review of the above contribution:

Round 1

- 2008-02-15: The members' chief editors submit comments and IDS analysis to the IRG chief editor.
- 2008-02-29: The IRG Chief Editor distributes consolidated comments to the members' chief editors.
- 2008-03-07: The members' chief editors submit confirmation to the IRG Chief Editor.
- 2008-03-14: The IRG Chief Editor distributes Urgent v1.1 direction to the IRG Technical Editor and members' chief editors.
- 2008-03-21: The IRG Technical Editor distributes CJK_Dv3 to members' chief editors.
- IRG#30: discussion on D set of Urgent v1.1.

2. CJK_D

2.1 process of CJK_Dv22 D set

The editors discussed all unification pairs with mark “N” or no mark which means no unanimous comments reached before this meeting (IRGN1353_CJK_D_V22-D_Unification.pdf and related excel file). It was agreed that all unification pairs without conclusion at this meeting should be removed from CJK_D for further study, these characters could be re submitted to future CJK extensions if unification issues resolved later.

The editors discussed all radical issues with mark “N” or no mark which means no unanimous comments reached before this meeting (IRGN1353_CJK_D_V22-D_Attributes.pdf and related excel file).

The editors agreed that all unification pairs and radical questioned characters with mark “Y”, which means unanimous comments reached

before this meeting, could be kept in CJK_D in principle. The IRG Chief Editor was required to double-check these characters and all other characters in CJK_Dv22 D set, which were not reviewed at the meeting, to insure the quality of CJK_D next version.

Note 1: 4 pairs of T source characters were temporally removed from CJK_D. Whether or not including them in CJK_D depends on the related discussion of Annex S of ISO/IEC 10646. (06985vsU+08413, 03665vsU+2AC34(TD-3B6D), 04501vsU+06f01, 07423vsU+087f3, ref: IRGN1385 TCA)

Note 2: G_CYY and G_ID characters of unification pairs with mark “N” or no mark were accepted to CJK_D in principle. China was required to provide convincing information (place names and pronunciations) of these characters by the next IRG meeting, or they should be removed from CJK_D. (ref: IRGN1384 item 1 China)

All G source characters with component “帶” were agreed to be kept in CJK_D for the reason that they are simplified and actually used. (ref: IRGN1384 item 3 China)

5 pairs of characters (05218vsU+753E, 02524vsU+21e1f, 00018vsU+53C0, 01070vsU+20D64, 03871vsU+2348B and U+23555) were temporally removed from CJK_D. Whether or not including them in CJK_D depends on the related discussion of Annex S of ISO/IEC 10646.

Note 3: All characters discussed at the meeting should not be merged to the latest M set to prevent being discussed again.

2.2 process of a few of ROK characters

The editors reviewed ROK proposal IRGN1378 and agreed to process the raised ROK characters in CJK_D as follows:

- Change of glyph for K5H00249 accepted.
- Change of source reference from K5H0dddd to K5-hhhh accepted.
- Change the glyph K5H00029, K5H00141 accepted but need to check for possible unification.
- K5H00029 should be added to next version of CJK_D because it is missing in the latest version by technical error.

2.3 67 missing characters

It is found that 67 characters removed from CJK_C balloting document are not included by CJK_D latest version. The editors agreed to add them to the D set of next version for further discussion.

2.4 other issues related to CJK_D

China was requested to prepare a document on problems of Kangxi representatives (康熙字典字頭字) and on possible wrong Kangxi glyphs for further discussion. TCA was invited to help China to find the counterparts of 3 raised compatibility characters. (ref: IRGN1384 item 2 China)

2.5 CJK_D review schedule

Round 1

- 2007-12-27: The IRG Chief Editor distributes double-checked CJK_Dv3 direction to members' chief editors.
- 2008-01-25: The members' chief editors submit confirmation to the IRG Chief Editor.
- 2008-02-15: The IRG Chief Editor distributes CJK_Dv3 direction to the IRG Technical Editor and members' chief editors.
- 2008-02-29: The IRG Technical Editor distributes CJK_Dv3 to members' chief editors.
- IRG#30: discussion on D set of CJK_Dv3

3. CJK_C

The editors reviewed related WG2 resolution M51.11 and required the IRG Technical Editor to update CJK_C index in time.

The editors reviewed and discussed IRGN1379 (TCA). The editors required the IRG Technical Editor to update CJK_C balloting document in time for 3 missing characters and re-sorting 14 characters. The editors requested IRG to submit a document to WG2 for this issue.

4. CJK_B multi-column code charts

The CJK_B multi-column code charts (IRGN1381) was distributed to IRG members during the meeting. The editors set the following schedule for CJK_B multi-column code charts review:

Round 1

- 2008-03-14: Members' chief editors submit review results to the IRG Chief Editor. (New fonts should be submitted if fonts correction requested.)
- 2008-03-28: The IRG Chief Editor distributes consolidated comments to members' chief editors.
- 2008-04-04: Members' chief editors submit confirmation to the IRG Chief Editor.
- 2008-04-11: The IRG Chief Editor distributes confirmed comments to the IRG Technical Editor.
- 2008-04-18: The IRG Technical Editor distributes CJK_B_MultiColumnV2 to members' chief editors.
- IRG#30: Final confirmation of CJK_B Multi-Column code charts.

5. 6 Japan characters in BMP and Compatibility Zone

The editors reviewed WG2 resolution M51.10 and WG2N3318. The editors were required to provide feedback to 1 character to be proposed to CJK Unified Ideographs Zone before the IRG plenary, no response would be considered confirmed. (No feedback so far.)

Appendix:

1. CJK_D discussions, recorded by Mr. Retarkgo Yan
2. Timetable (excel file)

附件 11

**IRG N1350 : Consolidated Comments
on Annex S Revision**

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2/IRG N1350

Date: 2007-11-05

Source:	Annex S Ad Hoc Group
Title:	CJK unification rules ad hoc
Status :	
Actions required	For discussion at IRG #29
Distribution:	Electronic
Medium :	

The Ad Hoc Group started the process of reviewing Annex S.

The ad hoc group notes WG2 resolution M51.36 "With reference to item 4 in document N3283, WG2 endorses the IRG activity to review and feedback on Annex S of ISO/IEC 10646: 2003, without impacting already unified ideographs in the standard, taking into consideration the FCD ballot progression which starts in 2008-03." In particular it is noted the above requirement not to 'impact already unified ideographs in the standard' and that it can only endorse suggestions the meet this requirement.

The review process included the circulation of a draft document of possible additions to Annex S, and the appendix A, B and C of this document are the various parts of that document taking into account comments from ad hoc members. Appendix D consists of suggestions regarding to clarifying Annex S by

making changes, rather than additions to Annex S.

The draft document was in html format, in response to feedback this version is in pdf format. Furthermore, with the exception of a suggested addition to S.1.5, references as to where additions might be put in Annex S have been removed.

The Chef Editor would like to thank all those who contributed to the discussion process whether ad hoc members or not.

Overview of the Appendixes

1) Possible additional examples (Appendixes A and B)

The ad hoc group started the process of considering what examples might be suggested for addition to Annex S.

A list of 130 possible examples of pairs/triplets of glyphs with the same abstract shape was circulated among the ad hoc members and no formal objections received about the choices, however comment was made on possible copyright issues of some of the images used, and therefore certain images have been replaced in the newer version of the list, appendix A.

A list of 158 examples of sets of glyphs with differences in abstract shape was circulated among the ad hoc members and formal objections expressing different points of view were received about one set, 古 vs 右. The original list contained 5 examples, already in the current Annex S which have now been removed, and 2 sets where for reference only, which have also been removed. The revised list, appendix B, has only 151 examples, which includes the 143 non-simplified vs simplified components in 1(i) below.

The two lists of possible examples still need to be discussed in full, in some cases additional information needs to be added, and for a few examples will need to be removed for more in depth research and discussion. The implications the examples still need to be considered carefully. Further it should be noted that the lists are not closed and suggestions of further examples are invited.

1.(i) Simplified Characters (Response to IRGN1302 and IRGN1308).

IRGN1302 notes two sets of examples.

The first set :-



Here it should be noted that Annex S itself in S.1.4 Examples of differences of abstract shape section S.1.4.3 includes 奂 and 奂 (U+5950 and U+5942) . It was therefore a case of over-unification to unify these characters.

The second set:-



However Annex S itself in section S.1.5 shows the right hand components 晉 and 晉 (U+6649 and U+664B) as having the same

abstract shape. It is therefore clear that in this case the unification rules have been applied correctly.

It is acknowledged that which are the standard simplified forms of characters may change with time, however the principles of Annex S can not change with time.

In IRGN1308 it was mistakenly suggested that all 146 pairs of standard simplified components in Appendix A are cases of different abstract shape, and that in Appendix B all the pairs of old vs new forms of characters are cases of the same abstract shape. This is not the case. Section S.3 includes a simplified vs non-simplified pair with the same abstract shape, 将 (U+5C06) and 將 (U+5C07). Furthermore Appendix B not only contains some of the pairs in the Appendix A list but also the pair 奂 and 奂 (U+5950 and U+5942), which is in fact an example of components of different abstract shape.

During the review process it was suggested that simplified characters be treated as a special case with their own section in the annex, however valid objections were made to this. It is noted that the annex itself does not treat simplified and non simplified forms as a special case, whilst most simplified variants have a different abstract shape to their non-simplified counter parts, this is not true for all simplified variants. To enable the work of unification of new characters to be carried out in an orderly manner it is suggested that additional examples of pairs of non-simplified vs simplified variants be added to the examples of same and different abstract shape in Annex S, and that these include, but not be restricted to, the 143 standard pairs not in the current version of Annex S .

It was therefore concluded that no group of variants be the simplified vs non-simplified, old vs new, or partially simplified

characters matched exactly the boundaries of Annex S, and that the unification process must be applied to all CJK unified ideographs.

Whilst due consideration can be given to adding examples to the annex, and on ideas regarding improving the presentation and clarity of Annex S, it is beyond the mandate of this ad hoc to suggest changing any of the already established unification patterns in Annex.

2) Possible additions (other than examples), clarifications and corrections to Annex S

2(i) The possible additional sections (Appendix C)

A number of possible additional paragraphs and sections were circulated for review. Corrections of grammar and spelling have been incorporated. Also, as detailed above, the suggestion not to have a special section for simplified characters. Regarding the section on 'Stability of unification rules' it was suggested that a sentence be removed, simple removal however would have distorted the section, rather the part has been rephrased bearing in mind the concerns expressed in the original objections.

No comments were received during the review process about 'An exception to S.1.4.2', however it is noted that in IRGN1353_CJK_D_V22-D_Unification.pdf where such cases are discussed there are differences of opinion about this concept.

2(ii) Possible clarifications and corrections (Appendix D)

The review process was also helpful in illustrating what could be clarified in Annex S and even some errors, such as the label 'most superior node' in Fig S.2.

附件 12

IRG N1397 : Annex S Review Ad Hoc Group Report

IRG N1397

Annex S Review Discussions

The following general points were agreed upon:

- It was generally agreed upon that some specific high-level components should not be unified (such as the radical or phonetic unit, which is a significant part of the character forms in question). See IRG N1350 Appendix B, Part 2 for examples. On the other hand, if the difference in the components is insignificant, then the characters may be unifiable. See IRG N1383 for additional notes on the discussions among members with regard to “significance” or lack thereof, “high-level,” and other conditions to consider for unification or dis-unification.
- Members must provide information and good reasons 1) if they want to unify characters for which the rules specify dis-unification; or 2) if they want to dis-unify characters for which the rules specify unification. Note that “information and good reasons” can include a wide variety of source references, such as readings, meanings, and other information that can help to distinguish them as separate characters (to dis-unify), or to assert that they are the same abstract character (to unify).
- Members agreed that the list in Annex S, Part S.3 needs thorough checking to determine whether they have the same abstract shape, and if necessary, to create separate lists to distinguish them.
- In view of the lack of time, members agreed to review documents related to the Annex S discussion, as follows, to include their appendixes: IRG N1350, IRG N1354, IRG N1371, IRG N1383, IRG N1387, IRG N1391, IRG N1395, and IRG N1396.
- By end of 2007: TCA to supply examples of unification and dis-unification for future discussion.
- By end of January 2008: Complete First Draft of a revised Annex S, with new and changed text, three new tables (unification examples, dis-unification examples, and errata), and a revised figure.
- By March 15, 2008: Members submit feedback, to include additional examples for the three tables (with comments) to John Knightley.
- By April 15, 2008: Complete Second Draft of a revised Annex S.
- By May 15, 2008: Members submit feedback to John Knightley.
- By June 2, 2008: Complete Third Draft of a revised Annex S for discussion at IRG #30.

IRGN1350 AppendixB non-unifiable

- The 143 pairs of standardized non-simplified (traditional) and simplified characters shown in “Jianhuazi Zongbiao,” and currently not in Annex S, were proposed in an attempt to make the example cases of dis-unification in Annex S more exhaustive, and in some cases, explicit.
- According to Japanese members, there is no systematic way to simplify characters in Japan, meaning that there is no official table or publication as there is in China.
- The change of the existing rule to accept both simplified and traditional forms of a character is still a pending discussion because some members have a strong objection to accepting any further submission of this kind.

The proposed pairs of possible dis-unification examples were discussed by members.

The examples with numbers 131 through 281 have been discussed and accepted in principle by members with the following exceptions:

- 137 Suggest to postpone because it is not a good example.
- 138 Not concluded, and postpone for further discussion.
- 152 Postpone for further discussion.
- 227 Accepted, but to correct the character on the left, i.e. the traditional form of the “執” (U+57F7).
- 275 Correct the glyph of the character on the left, but postpone to check for verification.
- 279 Postpone and check further.

IRGN1350 AppendixA unifiable Part02.pdf

The following unification examples 076 through 130 have been discussed and accepted in principle by members with the following exceptions:

- 078 add a note to the example:
“It is unifiable if the component is serving as ‘hand’; otherwise it is not unifiable.”
- 087 add U+56EA as one more example component in the list.
- 101 add U+5C02 as one more example component in the list.
- 112 pending as the case may be limited to certain characters only.
- 113 pending as K-column (4th char) of 5313 may be added to the case.
- 117 remove the “hand” component to make the reference example more clear.
- 122 remove the “hand” component to make the reference example more clear.
- 124 pending images rework due to images not clear

125 accepted but remove the note “However note 將 ≠ □ according to S5.2.3”.
129 Pending to check if it is possible using the characters in the Notes as examples
130 pending

IRGN1350 AppendixD errors.pdf

Error01 – Fig. S.2

- John Jenkins has produced a revised version of S.1.3.2, and redrew the diagram for Figure S.2 - the most superior node of a component.
- Members will review and give feedback on the document “AnnexS-Rewrite” (IRG N1396).

Error02 – S.3 Source code separation examples

The suggested revision to S.3 in Annex S is as below, and is also reflected in the document “AnnexS-Rewrite” (IRG N1396):

“The pairs (or triplets) of ideographs shown below are examples of disunification due to the source separation rule described in clause S.1.6. Inclusion in this list does not imply that these ideograph pairs (or triplets) should otherwise have been unified or disunified, and additionally does not imply that these pairs (or triplets) are cognates, though many of them clearly are cognate.”

IRGN1350 AppendixA unifiable Part01

Only part of the document was gone through with the following decisions arrived at in principle by members:

001 – pending

003 – non-cognate

008 – non-cognate component. Pending a redraw of the glyph

010 – postpone (the pair are non-cognate as characters but confused cognate as the component is sometimes interchangeable)

IRGN1371 AnnexSissues

IRG N1371 was gone through. The example of U+4E11 and U+4E12 was specifically discussed, and it was agreed upon that all examples in the document will go into pending for further discussion and review. The second example is proposed for dis-unification, and the third and fourth for unification, though there is also the issue of dis-unification when used stand-alone versus when used as a component.

Other examples discussed

- For the simplified version of “gate” 門 (U+9580) as used in China and Japan, it is agreed upon to unify the simplified versions of “gate” (U+95E8) as used in China and Japan for future submissions. These two simplified forms of “gate” will be accepted for unification as long as they are in complementary distribution. An example of a “two-dot” simplified form was also given as a topic of further discussion, shown below:

The image shows two simplified Chinese characters for 'gate'. The first character is a standard simplified form with a radical on the left and a 'one' on the right. The second character is a 'two-dot' simplified form, which has two dots above the radical on the left and a 'one' on the right.

- For China’s request to add “D-03030 and U+614C” as one more dis-unification example to Annex S (see IRG N1384), it needs more discussion to be certain.
- For TCA’s suggestion of the “D-06985 versus D-08413” pair (see IRG N1385), to ignore the radical on top and put as an example in Annex S for dis-unification, it has been agreed to move to pending.
- A table that uses the concepts of cognate/non-cognate and abstract/actual shape was proposed as an addition to Annex S, pending further discussion (IRG N1387).

附件 13

IRG N1377 : IRG#29 Resolution

ISO/IEC/JTC 1/SC 2/WG 2/IRG
Ideographic Rapporteur Group (IRG)

Resolutions of IRG Meeting #29

Source: IRG Meeting #29

Place: San Jose, California, the USA

Date: 2007-11-12 to 2007-11-16

IRG Meeting #29, attended by delegates from China, Hong Kong SAR, Japan, Republic of Korea (ROK), TCA, Vietnam, the USA and the Unicode Consortium, has made the following resolutions:

Resolution IRG M29.1: Future Meeting Schedules (N1355)

Unanimous

The IRG resolves to adopt the following meeting schedules:

IRG #30	Busan, Republic of Korea, 2008-06-09 ~13
IRG #31	Chongqing, China, 2008-11-17 ~ 21
IRG #32	Hong Kong SAR, June 2009 (tentative, to be confirmed after obtaining approval from authority)
IRG #33	IRG is seeking host, November 2009 (tentative)
IRG #34	IRG is seeking host, June 2010

Resolution IRG M29.2: Extension B Multi-column Code Charts (N1381 and N1388)

Unanimous

Action: All Member Bodies

The IRG accepts the CJK Editorial Group Report (N1388) on the schedule of review of Extension B multi-column code charts. The IRG notes the following timetable:

2008-03-14	Members' chief editors submit review results to the IRG Chief Editor. (New fonts should be submitted if font corrections are requested.)
2008-03-28	The IRG Chief Editor distributes consolidated comments to members' chief editors.
2008-04-04	Members' chief editors submit confirmation to the IRG Chief Editor.
2008-04-11	The IRG Chief Editor distributes confirmed comments to the IRG Technical Editor.
2008-04-18	The IRG Technical Editor distributes CJK_B_MultiColumnV2 to members' chief editors.
2008-06-02	Members' chief editors to submit feedback and comments to the IRG Chief Editor
IRG#30	Final confirmation of CJK_B multi-column code charts.

Resolution IRG M29.3: CJK Extension C Follow-up (WG2M51.11, N1347, N1348, N1351, N1352, N1379, N1388, N1392 and N1393)

Unanimous

Action: Technical Editor and Rapporteur

The IRG accepts the CJK Editorial Group Report (N1388) and N1379 related to CJK Ext C in response to WG2M51.11. The IRG instructs the Technical Editor to generate N1393 (CJK Ext C V11 with the mapping table).

The IRG further instructs the Rapporteur to submit N1392 (Ext C Correction Request to WG2) and N1393 (CJK Ext C V11 with the mapping table) to WG2 Project Editor.

Resolution IRG M29.4: CJK Extension D Project (N1353, N1353P1, N1353P2, N1353P3, N1353P4, N1353P5, N1378, N1384, N1385, N1388)

Unanimous

Action: All Member Bodies

The IRG accepts the CJK Editorial Group Report (N1388) on the development of CJK Ext D, in particular on the issues of unification and radical classification of D Set.

The IRG notes the following review schedule:

Round 1

2007-12-27	The IRG Chief Editor distributes draft direction on double-checked CJK_Dv2.2 to members' chief editors for confirmation.
2008-01-25	Members' chief editors submit confirmation to the IRG Chief Editor.
2008-02-15	The IRG Chief Editor distributes CJK_Dv3 direction to the IRG Technical Editor and members' chief editors.
2008-02-29	The IRG Technical Editor distributes CJK_Dv3 to members' chief editors.
IRG#30	Discussion on D Set of CJK_Dv3

Resolution IRG M29.5: Urgently Needed Characters (WG2M51.39, N1199, N1251, N1279, N1289, N1305, N1324, N1364, N1366, N1367, N1369, N1370, N1373, N1379, N1380, N1380A, N1380B, N1386, N1386P1, N1386P2, N1386P3, N1386P4 and N1388)

**China, Hong Kong SAR, Japan, TCA,
the USA, the Unicode Consortium, Vietnam-Acceptance
ROK-Against**

The IRG resolves to accept the Editorial Group Report (N1388) concerning urgently needed characters. The IRG resolves to accept characters which are stable and are of modern use. For these reasons, the submission from ROK and some characters in China's original submission are excluded. The IRG resolves to include submissions from China partially, Japan, TCA, and the Unicode Consortium.

The IRG resolves to process the Urgently Needed Character Set (N1380) (containing 238 characters not yet unified) in parallel with the Extension D project based on their individually agreed schedules.

The IRG set the following working schedule:

2008-02-15	Members' chief editors submit comments and IDS analysis to the IRG Chief Editor.
2008-02-29	The IRG Chief Editor distributes consolidated comments to members' chief editors.
2008-03-07	Members' chief editors submit confirmation to the IRG Chief Editor.
2008-03-14	The IRG Chief Editor distributes UNC v1.1 direction to the IRG Technical Editor and members' chief editors.
2008-03-21	The IRG Technical Editor distributes UNC v1.1 to members' chief editors.
IRG#30	Discussion on D set of Urgent v1.1.

Resolution IRG M29.6: Review of Annex S (WG2M51.36, N1350, N1350AP1, N1350AP2, N1350B, N1350C, N1350D, N1354, N1360, N1371, N1383, N1387, N1388 and N1391)

Unanimous

Action: All Members of the Ad Hoc Group

The IRG welcomes members and experts to join the Annex S Review Ad Hoc Group. Any interested parties may contact John Knightley directly.

The IRG resolves to accept the Annex S Review Ad Hoc Group Report (N1396).

The IRG notes the following review schedule:

2007-12-31	TCA to supply examples of unification and dis-unification in separate tables for future discussion
2008-01-31	To complete First Draft of a Revised Annex S, with new and changed text, three new tables (unification examples, dis-unification examples and errata), and a revised figure
2008-03-15	Members to submit feedback, including additional examples for the three tables (with comments) to John Knightley
2008-04-15	To complete Second Draft of a Revised Annex S
2008-05-15	Members to submit feedback to John Knightley
2008-06-02	To complete Third Draft of a Revised Annex S for discussion at IRG #30

Resolution IRG M29.7: Development of IRG Principles and Procedures (WG2M51.35, N1357, N1372 and N1382)

Unanimous

Action: All Members of the Review Group

The IRG resolves to accept the recommendation in WG2M51.35 for developing a set of principles and procedures to guide its work, with due regard for the current WG2 Principles and Procedures. The IRG also accepts N1382 with a view to achieving better recording and retrieval of editorial decisions.

The IRG resolves to set up a review group to draft a set of IRG Principles and Procedures. The membership includes:

John H Jenkins (Chief Drafter)
Yasuhiro Anan (Co-Chief Drafter)
Chen Zhuang (China)
Wang Xiaoming (China)
Retarkgo Yan (Hong Kong SAR)
Cheng Wai-hong, Peter (Hong Kong SAR)
Satoshi Yamamoto (Japan)
Kim Kyongsok (ROK)
Park Jong Woo (ROK)

Tseng, Shih-shyeng (TCA)
Wei, Lin-mei, Selina (TCA)
Ngo Trung Viet (Vietnam)
Lu Qin (Rapporteur)

The review schedule is as follows:

2007-12-21	First Draft ready from John H Jenkins
2008-01-31	Members to provide feedback
2008-02-29	Second Draft ready
2008-03-31	Members to provide feedback
2008-04-11	Third Draft ready for review by the IRG and WG2
2008-05-11	Members to provide feedback
2008-05-26	Fourth Draft ready
(two weeks before IRG#30)	

Resolution IRG M29.8: Old Hanzi (pre-Qin) Encoding (N1337, N1343, N1344, N1345 and N1346)

**Unanimous
Action: Old Hanzi Expert Group**

The IRG resolves to accept the Report of the Old Hanzi Expert Group (N1337) and notes its work plan as follows:

2008-03-28	To complete the Oracle Bone inscriptions of the 180 th to 260 th radicals in Shuowen and submit the results to Professor Li Guoying
2008-04-10	Selina Wei to finish the consolidation of all the submissions and email to all members of the Expert Group for review
2008-06-02	To submit the commented Oracle Bone inscriptions character table to the Rapporteur

Resolution IRG M29.9: Review of Japan's Submission to WG2 (WG2M51.10, N1347, N1347A and N1354)

**Unanimous
the USA--Abstention**

The IRG reviews the document from Japan (N1347) and accepts six characters in the proposal detailed in the CJK Editorial Group Report (N1388). The IRG instructs Japan national body to revise its proposal for submission to WG2 directly.

Resolution IRG M29.10: ROK's Request for Processing of the Remainder of Urgently Needed Characters Submission (N1394)

Unanimous

The IRG accepts the contribution from ROK in N1394 and resolves to ask all member bodies to put forward solutions for discussion at IRG#30.

Resolution IRG M29.11: Appreciation

By Acclamation

The IRG would like to express its sincere appreciation to the IRG#29 meeting host, Adobe Systems Incorporated and the Unicode Consortium. The IRG would also like to thank Mrs Magda Danish, Dr Ken Lunde and Mr Eric Muller for their excellent logistics, arrangements and hospitality.

The IRG also thanks Dr Ken Lunde, and Dr Richard Cook and Mr Tom Bishop, for their informative and interesting presentations on Ideographic Variation Sequence and Character Description Language respectively.

End of document