

出國報告(出國類別：其他)

出席 ISO/IEC JTC1/SC2/WG2/IRG 第 49 次會議出國報告

服務機關：經濟部標準檢驗局

姓名職稱：陳副局長玲慧、邱科長垂興

派赴國家：美國 加州聖荷西市

出國期間：106 年 10 月 15 日至 106 年 10 月 22 日

報告日期：106 年 12 月 11 日

摘要

ISO/IEC 10646 為全球統一碼字集(universal character set, UCS)之世界統一標準，收錄世界各國文字及符號之字形及編碼，制定之目的係為解決全球文字資訊交換、作業系統及應用軟體全球化等問題。

此行參與審查 IRG 之 2015 年度工作字集第 4 版、修訂 IRG 審字原則與程序之規章、更新可認同字元部件變體(UCV)及非可認同字元部件變體(NUCV)之 IRG 工作文件表列(IWDS)、討論表意文字變體資料庫(IVD)之認同規則、初步審查 IRG 之 2017 年度工作字集等，提出我國之技術貢獻，與各會員體之技術專家廣泛交流。

ISO/IEC 10646 為我國長期參與制定及維護之國際標準，後續 IRG 會議將持續討論相關表意文字及符號之擴編工作，我國應持續參與相關標準化活動，並與各會員體之技術專家維持良好互動及合作，進而維持我國中文資訊編碼對國際標準之影響力。

目次

| | |
|---------------|----|
| 壹、會議目的..... | 3 |
| 貳、會議簡介..... | 4 |
| 參、會議過程..... | 5 |
| 肆、心得與建議..... | 20 |
| 伍、重要名詞說明..... | 22 |
| 陸、活動照片..... | 26 |
| 柒、相關參考資料..... | 28 |
| 捌、附件..... | 29 |

壹、會議目的

ISO/IEC 10646 為全球統一碼字集(UCS)之世界統一標準，收錄世界各國文字與符號字形及編碼，制定之目的係為解決全球文字資訊交換、作業系統及應用軟體全球化等問題。

為推廣 CNS 11643 「中文標準交換碼」中文字集於國際間使用，以促使國際標準可符合我國資訊產業之需求，並展現我國對於國際標準之貢獻，我國多年來持續參與 ISO/IEC JTC1/SC2/WG2 及下屬表意文字書記組(IRG)之標準化活動，積極爭取將 CNS 11643 中文字集儘量全數納入 UCS。本局自 2001 年起即以專案計畫支持財團法人中文數位化技術推廣基金會之技術專家，聯合國內相關公(協)會及單位，以台北市電腦商業同業公會(Taipei Computer Association, TCA)之名義參與 WG2 及 IRG 之標準化活動。

爭取將 CNS 11643 中文字集納入 UCS 之中，使我國中文字編碼標準與國際標準同步一事，事關重大，若能將 CNS 11643 之中文字納入 UCS 中，則資訊設備使用者及資訊系統之開發者可逐漸使用更多之 CNS 11643 中文字，免除自行造字之困難，不致再發生顯示及列印出現亂碼之問題，對我國資訊產業及民眾存取資訊均有莫大助益。

此行參與 ISO/IEC JTC1/SC2/WG2/IRG 會議，參與審查 IRG 之 2015 年度工作字集第 4 版、修訂 IRG 審字原則與程序之規章、更新可認同字元部件變體(UCV)及非可認同字元部件變體(NUCV)之 IRG 工作文件表列(IWDS)、討論表意文字變體資料庫(IVD)之認同規則、初步審查 IRG 之 2017 年度工作字集等，提出我國之技術貢獻，與各會員體之技術專家廣泛交流。

ISO/IEC 10646 為我國長期參與制定及維護之國際標準，後續 IRG 會議將持續討論相關表意文字及符號之擴編工作，我國應持續參與相關標準化活動，並與各會員體之技術專家維持良好互動及合作，進而維持我國中文資訊編碼對國際標準之影響力。

貳、會議簡介

由於世界各國文字在字數、字形、使用方式及文化方面之差異，自人類使用電腦進行資訊處理以來，即面臨不同國家間資料交換之問題。國際間雖已有 ISO 646 資訊交換用 7 位元編碼字元集、ISO 2022 字元編碼結構及擴增技術等標準，可供各國據以制定其內碼或國家標準交換碼，互相進行資訊交換，但非固定長度之控制碼增添資料處理之困難度。此外，任何國家均不可能亦不必要將所有其它文字全數編碼在本國字集標準，在全球化時代中，處理本國以外之電子資訊成為各國需正視之問題。

為解決不同國家間電腦字元資訊交換之困難，自 1984 年起，國際標準化組織(ISO)與國際電工委員會(IEC)合組之第 1 聯合技術委員會下之第 2 分組委員會設立第 2 工作組(ISO/IEC JTC1/SC2/WG2)，積極推動 ISO/IEC 10646 之制定與修訂工作，即以集結全球通用之字符集，形成 1 個大聯集為目的，以滿足各國資訊交換之需求，提供全球語言文字與符號之表示、傳送、交換、處理、儲存、輸入及顯示之共同編碼標準，不但避免資源之浪費，並真正落實統一全球文字交換標準之理想。

由 ISO 與 IEC 合組之 JTC1 為專責資訊技術(information technology)之聯合技術委員會(joint technical committees)，其下共設立 20 個分組委員會(sub-committee, SC)，其中 SC2 負責編碼字元集(coded character sets)，而 SC2 下再設立第 2 工作組(Working Group 2, WG2)，專責全球統一碼字元集(Universal Coded Character sets)事務，即針對全球通用字集進行編碼，另鑑於中文字集存在因字意相同但筆劃或書寫方式不同而重複編碼之問題，對使用及維護造成困擾，因此於 WG2 下設立表意文字書記組(Ideograph Rapporteur Group, IRG)，專責中文字集之審查與整理。

參、會議過程

一、會議整體概述

ISO/IEC JTC1/SC2/WG2/IRG 第 49 次會議由美國之奧多比系統公司 (Adobe Systems Incorporated) 主辦，並提供會議場地及相關設施。

會議共舉行 5 天，討論議案包括各會員體之工作報告、審查 IRG 之 2015 年度工作字集第 4 版、修訂 IRG 審字原則與程序之規章、更新可認同字元部件變體(UCV)及非可認同字元部件變體(NUCV)之 IRG 工作文件表列(IWDS)、討論表意文字變體資料庫(IVD)之認同規則、初步審查 IRG 之 2017 年度工作字集等。會議整體概述如下：

| 日期 | 任務說明 | 備註 |
|-----------------|--|--------|
| 106 年 10 月 15 日 | 啟程前往美國 | 赴主辦國 |
| 106 年 10 月 16 日 | 1.會員體自 IRG 第 48 次會議後之工作報告 2.討論鬆綁表意文字變體資料庫(IVD) 認同規則之提案 3.討論香港補充字元集之橫向擴展及字元變更碼位提案 | 參與會議討論 |
| 106 年 10 月 17 日 | 1.討論中國大陸之字符更新提案 2.討論 IRG 審字原則與程序之規章(IRG PnP)修訂提案 3.討論中國大陸對於簡體字來源指定符之提案 | 參與會議討論 |
| 106 年 10 月 18 日 | 1.討論 IRG 工作文件序列(IWDS)更新提案 2.討論越南對於 ISO/IEC 10646 延伸 B 區中 V 單一來源字元之字符更新提案 3.討論中國大陸對於 G 來源字元及對應來源差異之提案 | 參與會議討論 |
| 106 年 10 月 19 日 | 1.討論 ISO/IEC 10646 延伸 F 區字元重複之提案 2.討論中國大陸對於更新部分 G 來源字元之提案 | 參與會議討論 |

| | | |
|-------------------------------------|--|---------|
| | 3.討論中國大陸對於新增 3 個表意文字 描述部件(IDC)之提案 4.初步審查 IRG 之 2017 年度工作字集 | |
| 106 年 10 月 20 日 | 確認本次會議決議事項 | 參與會議討論 |
| 106 年 10 月 21 日 ~106 年 10 月 22 日 | 自美返國 | 會議結束並返國 |

二、會議參與者名單

| 英文姓名 | 中文姓名 | 服務單位 |
|------------------|------|------------------|
| Chen, Zhuang | 陈 壮 | 中國電子技術標準化研究所 |
| Dai, Hong | 代 红 | 中國電子技術標準化研究所 |
| Zhou, Xiaowen | 周晓文 | 北京師範大學 |
| Tao, Yang | 陶 扬 | 中華字庫計畫執行成員 |
| Wei, Xiaohua | 魏巧华 | 中華字庫計畫執行成員 |
| Wu, Jian | 吴 健 | 中國科學院軟件研究所 |
| Tang, Yingmin | 唐英敏 | 北京大學 |
| Chen, Ken | 陈 恳 | 北京北大方正電子有限公司 |
| Huang Xuejun | 黄学钧 | 北京北大方正電子有限公司 |
| Lu, Qin | 陸 勤 | 香港理工大學 |
| KIM, Kyongsok | | 釜山大學 |
| Shin, Sang-hyun | 申相賢 | 高麗大學 |
| Cho, Sungduk | 趙成德 | 成均館大學 |
| Choi, Jeongdoh | 崔丁途 | 韓國國立國語院 |
| Toshiya Suzuki | 鈴木俊哉 | 廣島大學 |
| Wataru Takagi | 高木 涉 | 日立股份有限公司 |
| Wei, Lin Mei | 魏林梅 | 財團法人中文數位化技術推廣基金會 |
| Chen, Ling-hui | 陳玲慧 | 經濟部標準檢驗局 |
| Chiu, Chui-Hsing | 邱垂興 | 經濟部標準檢驗局 |
| Ken Lunde | 小林劍 | 奧多比系統股份有限公司 |
| Richard S. Cook | 曲理察 | 全球統一碼(Unicode)協會 |
| Deborah Anderson | | 個人 |
| Jaemin Chung | | 個人 |
| Lisa Moore | | 全球統一碼技術委員會(UTC) |
| Michel Suignard | | 全球統一碼(Unicode)協會 |
| Peter Constable | | 全球統一碼(Unicode)協會 |

| 英文姓名 | 中文姓名 | 服務單位 |
|---------------|------|-----------------------|
| Yinghui Zhang | 张映晖 | 中國社會科學院研究院 |
| Liang, Hai | 梁 海 | 個人 |
| Lee Collins | 康立論 | 越南喃字遺產保護會 |
| Wang, Yifan | 王一凡 | 日本佛典協會/一般財團法人人文學情報研究所 |

四、 會議過程摘要

(一) 106 年 10 月 16 日

1. 首先由 IRG 主席致詞，歡迎各會員體之代表齊聚美國加州聖荷西開會。復由與會代表逐一進行自我介紹。
2. 接續由 IRG 專案編輯報告前次會議決議事項及辦理情形。
3. 由 IRG 主席進行會議議程之確認，與會代表無異議通過。
4. 由各會員體報告自 IRG 第 48 次會議後之工作進展。

(1) 澳門工作報告

澳門未派代表參加本此 IRG 會議，由 IRG 主席代為宣讀書面工作報告，重點如下：

- ① 澳門特別行政區政府在香港理工大學協助下，於 2017 年 6 月成立澳門特別行政區資訊系統中文字元編碼統一架構，以作為資料交換編碼標準，增進澳門政府部門間資料交換之效率。
- ② 此統一架構首先在工作組與政府部門間運作，接續對所有政府部門公開，最後於對澳門民眾公開。此統一架構包含建立「澳門補充字元集」(MSCS)，補充大五碼(Big-5)與 HKSCS-2008 標準(香港補充字元集標準 2008 年版)。MSCS 第 1 版具備 166 個字元及符號，包含 23 個新增中文字元及 HKSCS-2016 中 1 個新增符號，滿足香港與澳門間電子資訊交換之需求。
- ③ ISO/IEC 10646 已包含 MSCS 中所有字元。在此架構公開後，將對 IRG 提出更新擴增字需求，並向全球統一碼(Unicode)協會申請表意文字變體序列(Ideographic Variation Sequences, IVS)註冊。

(2) 英國工作報告

由 Ken Lunde 代為宣讀書面報告，重點如下：

英國與來自美國、中國大陸、日本及俄羅斯之學術單位及技術專家討論後，向 IRG 提交 1,000 個字元作為 IRG 之 2017 年度工作字集(working set)。所提交之字元包含 243 個古漢字抄本表意文字、413 個方言用表意文字、170 個地方姓名用表意文字及 170 個其他表意文字。

(3) 我國工作報告

由財團法人中文數位化技術推廣基金會(以下簡稱中推會)魏林

梅秘書長進行報告，重點如下：

- ①為整合各國技術專家對於小篆字元編碼之意見，台北市電腦商業同業公會與中推會於 2017 年 8 月 28 日至 29 日共同舉辦特別工作組會議(ad-hoc task meeting)，邀請美國、日本及中國大陸與我國之技術專家進行研討，主要決議事項如下：
 - Ⓐ 說文解字小篆字元宜視為獨立字元，不隸屬中日韓非認同表意文字區(CJKUI)，另於新字元區(new block)獨立編碼。
 - Ⓑ 由於說文解字小篆字元獨立編碼，因此不適用 ISO/IEC 10646 中附錄 S 之 CJK 表意文字認同字處理規則，應另行制定處理規則。
 - Ⓒ 以藤花樹本所收錄之小篆字元為基礎，進行編碼。
- ② 針對 IRG 之 2017 年度工作字集，提出下列審查意見：
 - Ⓐ 部分字集未提供來源依據，且無對應之參考文獻。
 - Ⓑ 部分字集之來源依據不足。
 - Ⓒ 部分字元重複。

建議 IRG 會員體改善 IRG 之 2017 年度工作字集，於 IRG 第 49 次會議後重新提交。

此外我國對於提交 IRG 之 2017 年度工作字集補充說明如下：

- Ⓐ 我國提交 998 個擬納入 ISO/IEC 10646 中認同表意文字擴展區之新增字，作為 IRG 之 2017 年工作字集。其中包含 13 個化學字元及 985 個教育部異體字典中之表意文字。教育部異體字典係由 63 部字典中收錄超過 100,000 個中文字，為國內使用最為廣泛之線上字典。
 - Ⓑ 我國所提之新增字係反映我國之用字需求，期爭取納入 ISO/IEC 10646，以作為全球資訊相關廠商開發作業系統、應用軟體時可納入更多中文字，免除造字之困擾，可於各種資訊平台(包括手機、平板電腦、筆記型電腦、桌上型電腦等)正常顯示與列印中文字，避免產生亂碼，維持日常生活中各項與資訊交換之作業(文書處理、戶政、役政、國際貿易、工商登記等)得以順利運作。
 - ③ 經初步審查 IRG 之 2015 年度工作字集後，認為部分字元集尚有疑義，建議重新進行細部審查。
- (4) 日本佛典推廣團體(SAmganikikrtam Taisotripitakam, SAT)之工作

報告

- ① 協助審查 IRG 之 2015 年度工作字集。
- ② 提交 IRG 之 2017 年度工作字集。
- ③ 參與日本 SC2 委員會(對應於 ISO/IEC JTC1/SC2 之國內委員會)於 7 月 3 日召開之例會。
- ④ 於 7 月 14 日發表納入全球統一碼字集 10.0 版(Unicode 10.0)之字元集。

(5) 韓國工作報告

- ① 審查 1,475 個 IRG 之 2015 年度工作字集，提供相關審查意見。
- ② 針對 IRG 會員體審查韓國提交 IRG 之 2015 年度工作字集之審查意見撰寫回應資料。
- ③ 提交 IRG 之 2017 年度工作字集，包含 686 個漢字。

(6) 日本工作報告

- ① 派遣 3 名技術專家出席 6 月在韓國舉辦之 IRG 第 48 次會議、派遣 4 名技術專家出席 9 月在中國大陸舉辦之 WG2 第 66 次會議及 SC2 第 22 次會議。
- ② 對於 IRG 之 2015 年度工作字集之審查，由於部分日本專家反映審查時間不足，無法完成細部審查，建議修訂 IRG 審字原則與程序之規章，檢討審查時程與 IRG 會員體提交之新增字數。
- ③ 日本準備提交 1,645 個字元，建議納入 ISO/IEC 10646 之 CJK 認同表意文字擴展 F 區，目前正進行擬提交字元之國內查作業。

(7) 中國大陸工作報告

- ① 預定於 2018 年 5 月 21 日至 25 日於北京舉辦 IRG 第 50 次會議。
- ② 提交 969 個字元作為 IRG 之 2017 年度工作字集，包含 390 個壯字，其他則為地名用字、姓名用字、科學與技術用字等。中國大陸對於壯字補充說明如下：

- Ⓐ 壯字為壯族人民為書寫壯語而借用漢字之字形或結構所創造之文字。壯字最初創造之時代不明確，但已知 7 世紀即存在，最早記錄為唐代(公元 683 年)之石碑「六合堅固大宅頌」，及公元 697 年的智城碑，迄今已有 1,300 年

以上歷史。根據民間流傳的經文、詩詞等書面記錄、1989年編寫之「古壯字字典」收載大約 4,900 個正體字，正體字與異體字合計 1 萬餘字。雖自 1957 年以後，壯語出現應用羅馬字之壯文，但仍有人使用原始壯字。自中國大陸創造簡體字後，壯字中之漢字或漢字部件亦隨漢字之簡化而簡化。

⑤ 目前壯字之使用正逐漸式微，其應用範圍主要為山歌唱本及民間傳說，以及記錄戲曲劇本、藥方與家譜。

③ 收錄 1 百萬個漢字之「中華字庫」專案計畫，正針對其中 57 萬個漢字進行應用與推廣。計畫執行單位中國科學院軟體研究所已建立基於雲端運算技術之字型服務機制，任何應用程式及網頁可呼叫雲端應用程式介面 (Application Programming Interface, API) 即可取得中文字型資料，而不需在電腦中預先安裝字型，即可正常顯示、列印及進行資料交換。

④ 北京大學發布基於自動生成演算法所發展之中文手寫字型學習模式，僅需以 400 個手寫文字為樣本，即可將 CJK 區及擴展 A 區、B 區中所有字元自動產生類似手寫之字型，亦可同時將對應字型進行壓縮，僅需原始字型 25 % 之儲存空間，且失真程度極小。

(8) 美國工作報告

① 全球統一碼字集第 10 版業於 2017 年 6 月 20 日發布，收錄之總字數達 136,690 字，新增字共 8,518 字，包含 7,473 個 CJK 認同表意文字擴展 F 區之新增字、21 個 CJK 認同表意文字區之新增字、285 個日本變體假名、739 個中國大陸之女書文，CJK 認同表意文字數增至 87,882 字。全球統一碼字集第 10 版之核心規範將同時以 PDF 檔及紙本(共 2 冊)發布。

② 全球統一碼字集第 11 版預定於 2018 年 6 月發布。

(9) 香港工作報告

① 中華語言介面諮詢委員會於 2017 年 5 月發布最新版之香港補充字元集，名為 HKSCS-2016。與前一版 HKSCS-2008 相比，新版增加 23 個文字及 1 個符號，以滿足香港對於電子資訊交換之需求。

② 政府資訊科技總監辦公室及公務員事務局之法定語文事務部，共同審查 CJK 之 2015 工作字集，提供關於字型設計、部首、表意文字描述序列、筆劃數等提供意見，將提交 IRG 第 49 次會議討論。

2. 關於鬆綁 IVD 認同規則

(1) WG2 請 IRG 主席對於鬆綁全球統一碼技術規範 37 中 IVD 登錄規則研擬修訂文字，並提供合適之解釋及範例。由 IRG 主席與 Ken Lunde 博士共同草擬完成後，IRG 接獲 Henry 之回饋意見。

(2) 經技術專家群審查並修正文字內容，最終內容將送交全球統一碼技術委員會確認。

3. 關於香港補充字元集(HKSCS)之橫向擴展並更新字符更新並已納入 ISO /IEC 10646 國際編碼標準字集之字符，補充字源資料及字形

(1) 香港提案將 HKSCS-2016 中之 23 個字元及其更新之字符以橫向擴展方式補充字源資料及字形，經技術專家群審查並修正提案文件後，交由 WG2 確認。

(2) 橫向擴展(horizontal extension)係指對於已納入 ISO/IEC 10646 之字符，補充字源資料及字形。例如不同國家對於同一字但字形略有差異時，則以並列方式呈現，將該字形之來源地以 1 個英文字母表示，例如 G 代表中國大陸、H 代表香港、T 代表我國、J 代表日本、K 代表韓國、V 代表越南。例如：編碼為 4FAF 之「侯」



4. 關於香港補充字元集中 1 個字元變更碼位

香港提案將 H-8FA8 字元之碼位由 U+2F9B2 變更至 U+270F0，全球統一碼技術委員會(UTC)接受香港之提議，並提供 U+2F9B2 之 UIC 字符。經技術專家群審查並修正提案文件後，交由 WG2 確認。

(二)106 年 10 月 17 日

1. 關於中國大陸之字符更新

中國大陸提案將「通用規範漢字表」中 13 個字元及 11 個更新字符以橫向擴展之方式補充字源資料及字形，經技術專家群審查後，請中國大陸將 24 個已編碼之字符納入提案文件，將於下次 IRG 會議中討論。

2. IRG 審字原則與程序之規章(IRG PnP)

有技術專家表示簡體字來源並未於 IRG PnP 中定義，IRG 主席提議

於 IRG PnP 中新增簡體字來源之定義文字，經技術專家群討論後，確定文字內容。因受限於會議時間，對於命令參引(authoritative reference)之議題尚未討論。

3. 關於簡體字來源指定符

- (1) 編製 IRG 之 2015 年度工作字集時，簡體字來源在無實際使用依據之情況下如何處理之議題，例如提案文件之字符為字典掃描之影像，說明文字為「簡體字 B 之來源為繁體字 A」，IRG 主席提議對簡體字附加來源繁體字之指定符。
- (2) 美國技術專家不贊成此項提議，認為此提議在 20 年前或許可行，但時至今日許多簡體字之來源繁體字已個別編碼，僅適用於某些來源繁體字尚未被編碼之簡體字。
- (3) 日本技術專家則贊成此項提議，認為附加指定符之簡體字可視為來源繁體字之異體字。有技術專家認為此項提議可讓使用者可區別簡體字與來源繁體字之差異，但質疑是否所有相關之簡體字均存在對應之來源繁體字。主席表示，此提議係利用 2 個指定符，1 個用於指定簡體字之來源繁體字，另 1 個則用於指定來源繁體字對應之簡體字。
- (4) 中國大陸技術專家表示，中國大陸未來有更多簡體字來源之需求，少數需引用來源繁體字，僅某些廣泛使用之術語可能引用來源繁體字，因此不贊成附加指定符之方式。
- (5) 由於經討論後對仍無法形成共識，IRG 主席表示撤銷此項提案。

(三)106 年 10 月 18 日

1. 關於 IWDS 更新

網頁之鏈結將更新。其他維護議題將待相關意見提出後再行討論。

2. 關於 ISO/IEC 10646 延伸 B 區中 V 單一來源字元之字符更新

越南提案更新 2 個 ISO/IEC 10646 延伸 B 區中 V 單一來源字元之字符，雖然某些技術專家表示此 2 個字符已被使用，且字符之變更涉及語意之改變，但考量此變更尚不致對既有之文書造成負面衝擊，IRG 同意此提案。

3. 關於 3 個 G 來源字元及對應來源之差異

- (1) U+2B1C9 之代表字符與其來源資訊不同，雖然正確之字符已在不同區段編碼，但發現目前 U+2B1C9 之字符顯示為另 1 個來源，因此需變更來源資訊。另 U+2A931 之代表字符亦與其來源資訊不同，由於正確之字符尚未編碼，因此 U+2A931 來源資訊

予以更新即可。

- (2) 此外 U+2AC00 之代表字符亦發現與其來源資訊不同，由於正確之字符尚未編碼，但與現行之代表字符不同，且無法與正確之字符予以認同，因此 IRG 請中國大陸重新研究後，再重新提案。

(四)106 年 10 月 19 日

1. 關於 ISO/IEC 10646 延伸 F 區之字元重複

技術專家發現 ISO/IEC 10646 延伸 F 區中 U+2DE17 字元與延伸 B 區中 U+247EC 字元重複，由於延伸 F 區收字時已針對字符進行詳細審查，但韓國不慎將錯誤之電腦輪廓字體(true type)之字型檔提交 WG2 專案編輯以致發生重複字之問題，IRG 已請韓國將正確之字形檔重新提交 WG2 專案編輯處理。

2. 關於更新部分 G 來源字元

中國大陸提案將 1 個「宋本玉篇」之來源字變更字源資料，並將 66 個「北大方正汉字内码字典」之來源字以橫向擴展之方式新增字源資料及字形，經技術專家群審查後認為尚有疑義，請中國大陸針對該提案重新檢討。

3. 關於新增 3 個 IDC 之提案

中國大陸提出新增 3 個 IDC 以描述某些字元之議案，雖然日本技術專家表示許多對映之字符已編碼，但未強烈反對此提案，因此將由 UTC 將再討論此提案，中國大陸將針對 3 個 IDC 進行造字，造字檔將提供 WG2 專案編輯進行後續作業。

4. 初步審查 IRG 之 2017 年度工作字集

- (1) IRG 之 2017 年度工作字集共收錄 5,160 個字元。提交情形如下：

- ① 中國大陸提交 969 個字元。
- ② 韓國提交 686 個字元。
- ③ 日本佛典協會提 323 個字元。
- ④ 我國提交 998 個字元。
- ⑥ 英國提交 1,000 個字元。
- ⑦ 全球統一碼技術委員會提交 193 個字元。
- ⑧ 越南提交 991 個字元。

- (2) 各會員體所提交字集經初步審查後，普遍存在表意文字描述序列、來源證明之資訊不完整，以及發現重複字之問題，仍有改善之空間，爰建請各會員體重新檢討，刪除重複字並補足相關資訊，重新提交修正後之字集，以利重新審查。

- (3) 各會員體應依下列原則重新提交修正後之字集：
- ① 修正後字集中不得包含原字集所無之字元，即不得藉機(夾帶)新增字元。
 - ② 刪除或修正字元時應明確說明。
 - ③ 應嚴格遵守 IRG 審字原則與程序之規章中發現 5 % 錯誤即退件之原則。
 - ④ 建議採用表意文字描述序列(IDS)表示 CJK 補充部件。
- (4) 請各會員體之主任編輯依下列時程辦理：
- ① 2018 年 3 月 16 日：請各會員體之主任編輯將更新後之資料提交 IRG 主席及 IRG 主任編輯。
 - ② 2018 年 3 月 30 日：IRG 主任編輯分發 IRG 之 2017 年度工作字集 0.1 版。
 - ③ 2018 年 4 月 20 日：IRG 會員體之主任編輯及技術專家提供交互審查意見。
 - ④ 2018 年 5 月 4 日：IRG 會員體之主任編輯針對交互審查意見提供回應。

5. 關於其他事項

我國之提案因會議時間受限，於下次 IRG 會議中審查。韓國提送橫向擴展之提案，因資料提送時間研務，亦於下次 IRG 會議中審查。

(五) 106 年 10 月 20 日

本日之會議係針對本次會議討論後所形成之各項共識，彙整成為決議事項，並逐一進行討論及確認。

1. 主席宣布，IRG 第 49 次會議經來自中國大陸、香港、韓國、美國、全球統一碼協會(Unicode)、我國、日本佛典協會、越南等會員體代表之技術專家，以及以個人名義參與之技術專家共同討論後，決議如下：

(1) 決議事項 1：後續會議之規劃

- ① IRG 第 50 次會議：預定 2018 年 5 月 21 日至 25 日於中國大陸北京市舉辦。
- ② IRG 第 51 次會議：預定 2018 年 10 月 22 日至 26 日於越南河內市舉辦。
- ③ IRG 第 52 次會議：暫定 2019 年 5 月 13 日至 17 日於香港舉辦。
- ④ IRG 第 53 次會議：暫定 2019 年 10 月 21 日至 25 日於中國

大陸成都市或昆明市舉辦。

⑤ IRG 第 54 次會議：舉辦時間與地點尚在研議中。

(2) 決議事項 2：審查 IRG 之 2015 年度工作字集第 4 版

IRG 討論各會員體代表之技術專家針對 IRG 之 2015 年度工作字集 4.0 版所提供之意見，並通過編輯報告。IRG 同意展開 IRG 之 2015 年度工作字集 5.0 版之編製作業，請 IRG 會員體將字集提交 ISO/IEC 10646 之專案編輯(project editor)，由專案編輯彙整為未來納入 ISO/IEC 10646 之延伸 G 區碼表後，進行下一輪之審查作業。審查作業期程規劃如下：

① 2017 年 10 月 27 日：由 IRG 主任編輯(chief editor)分發討論紀錄。

② 2017 年 11 月 3 日：IRG 會員體之主任編輯向 IRG 主任編輯提供交互審查意見，另將更新之基本多語文字面(basic multilingual plane, BMP)字型檔提交 IRG 主任編輯，並將電腦輪廓字體(true type)之字型檔提交 ISO/IEC 10646 之專案編輯。

③ 2017 年 12 月 1 日：由 IRG 主任編輯分發 IRG 之 2015 年度工作字集 5.0 版。

④ 2018 年 2 月 24 日：IRG 會員體之主任編輯及技術專家向 IRG 主任編輯提供交互審查意見。

⑤ 2018 年 3 月 16 日：IRG 主任編輯分發交互審查意見。

⑥ 2018 年 4 月 1 日：IRG 主席向 IRG 會員體之主任編輯分發草案之投票意見。

⑦ 2018 年 5 月 4 日：IRG 會員體之主任編輯針對交互審查意見及草案之投票意見向 IRG 主席提出回應。

(3) 決議事項 3：修訂 IRG 審字原則與程序之規章

IRG 針對各會員體回饋之意見，討論 IRG 審字原則與程序之規章第 10 版修訂草案後，完成 IRG 審字原則與程序之規章第 10 版之修訂，其中包含對於簡化字元之定義等修訂內容。IRG 主席將更新 IRG 網站中對於新版資料之鏈結，並將經確認後之 IRG 審字原則與程序之規章第 10 版提交 WG2 備查。

(4) 決議事項 4：更新可認同字元部件變體(UCV)及非可認同字元部件變體(NUCV)之 IRG 工作文件表列(IWDS)

IRG 決議分別更新本次會議已確認之可認同字元部件變體及非

可認同字元部件變體等 2 種 IRG 工作文件表列。IRG 另決議通過說明文件及韓國正規字元文件等 2 種 IRG 工作文件表列。IRG 主席將更新 IRG 網站中對於上述資料之鏈結。

- (5) 決議事項 5：討論表意文字變體資料庫之認同規則
IRG 已針對 WG2 之意見討論完畢，並將建議內容綜整，由 Ken Lunde 博士向全球統一碼技術委員會(UTC)回饋建議內容。
- (6) 決議事項 6：審查越南提交之橫向擴展字集
IRG 審查各會員體針對越南提交橫向擴展字集之意見後，請越南依各會員體之建議事項修正提案文件，並重提修正後之字集。
- (7) 決議事項 7：審查越南提交變更來源之字集
IRG 審查各會員體針對越南提交 10 個變更來源之字集後，同意變更為 V 來源字集。請越南依各會員體之建議事項修正提案文件，並重提修正後之字集，將檔案提交 ISO/IEC 10646 專案編輯及 IRG 主席。
- (8) 決議事項 8：審查香港提交之橫向擴展字集、字符更新及字元對映資訊
IRG 審查香港提出之橫向擴展字集、字符更新、字元對映資訊及更新 H 來源字符後，接受香港之提案。請香港將字元對映表及重新對映之資訊提交 ISO/IEC 10646 專案編輯。請 IRG 主席將 H 來源字符之字形檔提交 ISO/IEC 10646 專案編輯。
- (9) 決議事項 9：審查英國提交來源參考字集更新
IRG 審查英國提交來源參考字集更新後，同意將 IRG 之 2015 年工作字集中原為 UTC 來源字集轉為 UK 來源字集。
- (10) 決議事項 10：審查中國大陸提交 G 來源字集
IRG 審查中國大陸提交 2 個變更來源字集更新後，同意中國大陸之更新提案。
- (11) 決議事項 11：審查 IRG 之 2017 年度工作字集
IRG 審查各會員體提交之 2017 年度工作字集，以及交互審查之意見後，確認仍有不完整之處，請各會員體重新檢討，提出修正後 IRG 之 2017 年度工作字集 0.1 版，進行下一輪審查作業。請各會員體之主任編輯依下列時程辦理：
 - ① 2018 年 3 月 16 日：請各會員體之主任編輯將更新後之資料提交 IRG 主席及 IRG 主任編輯。
 - ② 2018 年 3 月 30 日：IRG 主任編輯分發 IRG 之 2017 年度工

作字集 0.1 版。

③ 2018 年 4 月 20 日：IRG 會員體之主任編輯及技術專家提供交互審查意見。

④ 2018 年 5 月 4 日：IRG 會員體之主任編輯針對交互審查意見提供回應。

(12) 決議事項 12：審查中國大陸提交表意文字變體字元

IRG 審查中國大陸提案新增 3 個表意文字變體字元後，同意中國大陸之新增字元需求。請中國大陸於 2017 年 10 月 20 日將字型檔提交 IRG 主席。並請 IRG 主席將字型檔轉交 WG2 及 ISO/IEC 10646 專案編輯。請 Ken Lunde 博士再將字型檔轉交 UTC。

2. IRG 對於 IRG 第 49 次會議之主辦單位、全球統一碼協會表達誠摯感謝，特別感謝 Ken Lunde 博士對於本次會議各項安排之貢獻。對於全球統一碼協會為與會之技術專家提供 2 冊全球統一碼第 10 版紙本書籍，IRG 亦表達感謝之意。

肆、心得與建議

一、心得

- (一) 物聯網時代來臨，日常生活之各種電機、電子、資訊及通信設備均可連接上網，而連接網路之設備需內建作業系統及軟體，資訊之提供均須仰賴文字，而 ISO/IEC 10646 為全球之資訊業者提供文字與符號之統一標準，儼然成為資訊交換之平台，以作為產品開發之依據，對於資訊交換媒介之需求日益擴大，ISO/IEC 10646 之重要性不言可喻。ISO/IEC 10646 後續將持續進行相關表意文字及符號之擴編工作，而 IRG 會議為我國少數可直接參與國際標準制定之管道，我國應持續參與該工作組之標準化活動，並與各國技術專家維持良好合作與互動關係，以擴大我國中文資訊編碼對國際標準之影響力。
- (二) 我國在漢字應用與推行方面，分由不同機關執行。教育部負責文獻蒐集及字形來源之鑑定，國家發展委員會負責全字庫之字形收錄及暫編碼之維護，文化部則負責文字之應用與推廣，移民署及各地戶政、役政機關負責接受新字之申請及初審，本局則負責國家標準制定及建立對國際標準提案之管道。在多年之努力下，本局已完成中文碼相關國家標準與國際標準之調和，並維持參與國際標準制定之管道，為國內資訊及文創產業在國際場域發聲，將我國特有漢字爭取納入國際標準，帶動產業發展，未來亦應保持與國內各機關間之協調及合作，將國內之需求適時反映至國際場域，為我國爭取最大利益。
- (三) 據參與 IRG 會議技術專家之經驗，由提案、正式編碼、納入國際標準並發布，最長需耗費 10 年，因此長期參與極為重要。為使我國之提案能獲得國際標準採納，除需精進專業能力並持續提出技術貢獻，以獲得國際肯定外，熟悉會議程序、掌握談判策略並具備溝通與協調能力亦極為重要，因此本局應持續支持及培養參與國際會議之專業人才。
- (四) 中國大陸積極將瀕臨失傳之古文字，爭取納入國際標準中，103 年提交女書文之提案，104 年提交水書文之提案，105 年提交西南古文之提案，今(106)年則提交壯族文字之提案，由此可見 ISO/IEC 10646 從以往僅針對解決資訊交換問題，滿足字形可跨不同載具及平台正常顯示之基本需求外，更進一步肩負文化保存與傳承之責任。我國爭取將特有漢字納入國際標準之提案目的，除文化保存之外，更具備協助國內文創產業發展之意義。例如小篆、甲骨文等古漢字已應用於影視節目、產品設計、圖文出版及遊戲開發等領域，除可發揚我國固有文化外，亦可創造商品化價值。未來我國應持續爭取將更多特有漢字納入國際標準。

二、建議

- (一) 我國自參與 ISO/IEC/JTC1/SC2/WG2/IRG 之會議以來，皆以台北市電腦商業同業公會(TCA)之名義出席會議，而 TCA 已多年未派員出席，但中推會與本局仍以 TCA 名義出席似乎不夠名正言順，因此可考量能否改以其他名義或個人名義出席。由於各會員體對於 TCA 之代表性已根深蒂固，以其他名義或個人名義出席會議能否獲得各會員體認同，宜審慎評估，擬先由中推會與 IRG 主席協商，並徵詢各會員體之看法後再進一步研議。
- (二) 早年財團法人資訊工業策進會、中央研究院計算機中心、中央研究院歷史語言研究所、行政院主計處電子處理資料中心(中文資訊處理業務已移撥國家發展委員會)、國立臺灣師範大學國文系、台灣國際商業機器(IBM)股份有限公司等機關/團體曾派員出席 IRG 會議，但近年來僅本局及中推會仍持續派員出席會議，鑑於日常生活許多事務、業務與中文資訊用字息息相關，及早掌握國際標準發展趨勢極為重要，後續會議可嘗試邀請相關機關/團體一同出席。
- (三) 我國曾於 97 年間爭取 98 年度 IRG 會議在我國主辦，復因中國大陸代表團表示無法赴臺出席會議，爰改由越南主辦。由於我國參與 IRG 會議多年，專業與貢獻深獲各會員體肯定，應可再嘗試爭取由我國主辦 IRG 會議，以提升我國之國際能見度，擬先由中推會與 IRG 主席協商，並徵詢各會員體之看法後再進一步研議。

伍、重要名詞說明

- [1] 國際標準化組織與國際電工委員會合組第 1 聯合技術委員會下第 2 分組委員會之第 2 工作組(ISO/IEC JTC1/SC2/WG2)

ISO(International Organization for Standardization)為國際標準化組織，IEC(International Electrotechnical Commission)為國際電工委員會，兩者均為國際標準之制定機構，而 JTC(Joint Technical Committees)為兩者合組之聯合技術委員會，其中 JTC1 為專責資訊技術(information technology)之聯合技術委員會，其下共設立 20 個分組委員會 (Sub-committee, SC)，其中 SC2 負責編碼字元集(coded character sets)，而 SC2 下再設立第 2 工作組(Working Group 2, WG2)，專責廣用編碼字元集(universal coded character sets)事務，即針對字形進行編碼。

- [2] 表意文字書記組(Ideograph Rapporteur Group, IRG)

係為 ISO 於 1993 年在 SC2/WG2 之下成立表意文字書記組，專司「中日韓認同表意文字」區(CJK Unified Ideographs)之整理工作。

中日韓漢字最初進行編碼時，因各國漢字形體不盡相同，須先進行認同(unify)整理工作，SC2/WG2 因此邀集相關各國指派專家組成 CJK/JRG(中日韓聯合研究工作組，即 IRG 前身)，進行字集之總整理。1993 年公布之 ISO/IEC 10646-1 中，已將我國、中國大陸、日本及南韓所使用之漢字整合為獨立之中日韓認同表意文字區，並予以編碼。

在「中日韓認同表意文字」區中，因空間有限，未能將亞洲各國所使用之表意文字全部納入，難以滿足各國實際應用之需求，因此各國咸認表意文字字集擴增之必要性。因此於 1993 年成立 IRG，專司「中日韓認同表意文字」區的內部/水平擴充(international/horizontal supplementation)及外部/垂直擴充(external/vertical extension)之整理工作。參與者包括中國大陸、我國、日本、韓國、香港、美國、越南及新加坡等。

IRG 成立以來，依循每年 2 次、各國輪流舉辦之慣例，討論 ISO/IEC 10646 表意文字字集之擴充原則，並就各國提出之字集進行認同及整理。

- [3] 全球統一碼(unicode)

實現全球語言文字與符號之表示、傳送、交換、處理、儲存、輸入及顯示之共同編碼系統。

由於各國文字在字數、字形、使用方式及文化方面之差異，自人類使用電腦進行資訊處理以來，即面臨不同國家間資料交換之困擾。國際間雖已有 ISO 646、ISO 2022 等編碼規範，可供各國據以制定其內碼或國家標準交換碼，

並互相交換；但非固定長度之控制碼增添資料處理之困難度。此外，任何國家均不可能亦不必要將所有其它文字全數編碼在本國字集標準，在全球化時代中，處理本國以外之電子資訊成為各國需正視之問題。

為解決不同國家間電腦字元資訊交換之困難，自 1984 年起，全球陸續出現 2 個組織，期發展出全球可以共用之編碼字集。其中之一為 ISO/IEC JTC1/SC2/WG2，其推動制定之 ISO/IEC 10646，即以集結全球通用之字符集，形成 1 個大聯集為目的，以滿足各國資訊交換之需求。另外一個組織則為全球統一碼協會(Unicode Consortium)，亦設法採用新觀念及架構，設計出適用全球之統一碼(Universal Code；簡稱 Unicode)。此 2 個組織之工作及方向一開始各自發展，但最後終於殊途同歸，整合為 ISO/IEC 10646/Unicode 標準，提供全球語言文字與符號之表示、傳送、交換、處理、儲存、輸入及顯示的共同編碼標準，不但避免資源之浪費，並真正落實統一全球文字交換標準之理想。

[4] 字元(character)

用以組織、控制或表示資料之元件組的構件。

[5] 字(元)集(character set)

由字元組成之集合。

[6] 字型(font)

具相同基本設計字符影像之集合。

[7] 字符(glyph)

可辨識之抽象圖形符號，與任何特定的設計無關。

[8] 字形(glyph shape)

字符表示法中的 1 組資訊，用以定義表示字符之形狀。

[9] CJK 表意文字(CJK Ideograph)

係指中日韓表意文字。我國、中國大陸、南韓與日本均使用漢字，以編碼之觀點，不應各自編碼，造成一字多碼之現象，與 ISO/IEC 10646 一字一碼之編碼基本原則抵觸。我國及中國大陸之代表即共同提出漢字集(Han Character Set, HCS)之構想，獲得到美國及其它國家代表支持，於在 1993 年公布之 ISO/IEC 10646-1 中，已將我國、中國大陸、日本及南韓所使用之漢字整合為獨立之中日韓認同表意文字區，並予以編碼，使得亞洲幾個主要使用漢字之國家，得以有圓滿之解決方案。

[10] 認同(unified)

即字義相同但字形不同之 2 個字元，換言之為書寫方式不同，則此 2 個字元認定為同一字，不再重複編碼，避免形成 1 字 2 碼。例如：編碼為 4E5F 之

「也」字，其中第一筆「一」香港(H 字源，字下方編碼之首個英文代號，字源係該會員體所提出之字符來源)及我國(T 字源)字形之書寫方式均無倒勾，中國大陸(G 字源)、日本(J 字源)、韓國(K 字源)、越南(V 字源)字形之書寫方式為「𠄎」，均有倒勾，但因字義均相同，故視為同一字，編碼均為 4E5F，則可避免重複編碼。



[11] 表意文字變體序列(ideographic variation sequence, IVS)

變體往往造成編碼方面之困難，許多字意相同但筆劃或書寫方式不同之字，僅存在些微差異即視為不同字，重複編碼之結果將造成一字多碼之問題，使用時易發生混淆且不利於字集後續之管理與維護。為避免重複編碼之問題，IRG 會議絕大部分之時間均針對變體依據中日韓表意文字認同規則，討論能否認定為同一字，儘量將字意相同但筆劃或書寫方式不同之字予以認同，將認同字列於 ISO/IEC 10646 之中日韓認同表意文字區，僅編 1 組字碼，即多字共用同一碼。以甲骨文為例，筆劃接近之變體可能在字義方面不同，因此不同變體均應賦予 1 組字碼，此時即不適用認同規則，為避免違反不得一字多碼之原則，可採用表意文字變體序列之技術。

此技術係由全球統一碼協會(Unicode Consortium)為漢字所建立之字形定義技術，其基本原理係將 1 個字符之基本字形用於該編碼字，例如「骨」字，其編碼為 <U+9AA8>，而與之認同字形「𠄎」字則可依基本字形之編碼 <U+9AA8> 另外加 1 組變體選擇符(variation selector)<U+E0100>，由兩者定義 1 個組合字序列<U+9AA8> <U+E0100>。採用表意文字變體序列時，變體仍以基本字形之編碼為準，符合 1 字 1 碼之原則，搭配可支援表意文字變體序列之軟體(目前奧多比系統公司旗下之 Acrobat 等軟體已支援該技術)，則對於「骨」字或「𠄎」字均可正確顯示，對於不支援表意文字變體序列之軟體，由於無法識別變體選擇符，僅能識別基本字形之編碼，因此僅能顯示「骨」字而無法顯示「𠄎」字，但不致顯示亂碼，反觀若不採用表意文字變體序列之技術，「骨」字可正確顯示，但「𠄎」字將顯示亂碼。

可用之變體選擇符由<U+E0100>至<U+E1EF>計 240 組，即每個表意文字之字符最多可關聯 240 個變體形。以表意文字變體序列所定義之變體，向全球統一碼協會申請註冊並收錄於表意文字變體資料庫(ideographic variation database, IVD)後，可為變體之資訊交換提供可符合 ISO/IEC 10646 編碼架構之方案。表意文字變體序列及資料庫對於以往在漢字認同規則下無法編碼之變體，透過由已編碼之漢字與變體選擇符所組成之變體序列，可與已編碼之

漢字相互關聯並取得編碼，再收錄於表意文字變體資料庫後，即可解決古漢字編碼之問題。

[12] 表義文字描述序列(ideographic description sequence, IDS)

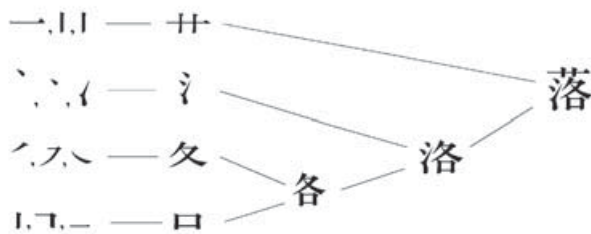
漢字字形之結構組成描述，即將字形拆解為不同基本部件(component)之組成描述。

[13] IRG 工作文件表列(IRG working document series, IWDS)

IRG 為進行認同字審查之作業，持續維護表意文字之文件集，用以補充 ISO/IEC 10646 附錄 S 「CJK 表意文字認同及安排之程序」中對應之範例，俾使 CJK 認同表意文字列表保持在最新之狀態。

[14] 可認同字元部件變體(unifiable component variations, UCV)

針對可認同字元，將字形拆解為基本部件之結構分析。例如「落」字，字形可拆分為上半部與下半部，上半部拆解之基本部件為「艹」，下半部拆解之基本部件為「洛」。下半部之基本部件「洛」可拆分為左半部及右半部，左半部拆解之基本部件為「氵」，右半部拆解之基本部件為「各」。右半部之基本部件「各」可再拆分為上半部與下半部，上半部拆解之基本部件為「夂」，下半部拆解之基本部件為「口」。各基本部件再進一步拆解則為書寫之筆劃，原則上字形拆解至基本部件已無法再進一步解構，即完成其結構分析，拆解為筆劃係便於計算字形之筆劃數。UCV 描述之範例如下：



以文件化表示方法如下，以「花」字為例：

花 = 艹 匕
 = 艹 匕 匕

陸、活動照片



會議討論情形



會議討論情形



團體照

柒、相關參考資料

- [1] CNS 5205 資訊技術-資訊交換用七位元碼字元集(ISO/IEC 646)
- [2] CNS 7654 資訊技術-字元碼結構及延伸技術(ISO/IEC 2022)
- [3] CNS 11643 中文標準交換碼
- [3] CNS 14649 資訊技術-廣用編碼字元集(UCS) (ISO/IEC 10646)
- [4] 全字庫網站(<http://www.cns11643.gov.tw/AIDB/welcome.do>)
- [5] ISO/IEC JTC1/SC2/WG2/IRG 第 49 次會議之開會資料
(<http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg49/IRG49.htm>)

捌、附件

- 1.ISO/IEC JTC1/SC2/WG2/IRG 第 49 次會議之議程
- 2.我國提出技術貢獻之相關資料
- 3.ISO/IEC JTC1/SC2/WG2/IRG 第 49 次會議之會議紀錄