

「Unicode 異體字對應表初稿」說明

2009/6/30

一、編訂目的

書目紀錄內容常因編目員輸入或出版品印製時，對於正異體字形的選用不同（例如：「峰」與「峯」），而影響相關書目資料的查獲率。「Unicode 工作小組」在「CCCII_Unicode 字碼對照表」的基礎上，經過多次會議討論，完成「Unicode 異體字對應表」（以下簡稱「對應表」）初稿之編訂，提供圖書館及自動化系統廠商下載運用，期望透過正異體(漢)字之對應，能讓使用者同時查到相對應的正體、簡體、異體字形資料，以提升圖書館目錄系統書目的查獲率。

二、編訂方式

本對應表先依據 CCCII 的正異體字碼架構擷取對應的字形及字碼，再逐一查核下列參考資料，無法在下列任一資料找到對應關係者，不予收錄：

- (一) 教育部異體字字典 (<http://140.111.1.40/main.htm>)
- (二) 遠東.漢語大字典之「異體字表」(遠東圖書公司，1991)
- (三) 實用繁簡體字手冊(臺灣商務印書館，2003)
- (四) 日本文化廳常用漢字一覽表

三、分級對應方式

為提高對應表運用之彈性，本表採分級對應方式，未來各圖書館或系統廠商可參考本表分級方式，並視實際運用情形及需求，斟酌調整對應範圍。目前分級原則如下：

(一)「第 1 層」

- 1、所有字形本身必須為繁體字，或臺灣地區通用的簡體字；
- 2、字義必須相等或互通，也必須能交互應用；
- 3、不論字形差異之多寡，亦不論是否有 Big5 碼，只要符合上述 1 及 2 之情形，皆與收錄；
- 4、對應的字組中，字義雖部份相通，但若有某些無法互通的用法，則不收錄於此層級，例如：「伙」雖為「夥」的簡體字，但「伙夫」不可作「夥夫」，因此，不予收錄於第 1 層；

5、字義不同，字形相似，時有誤用情形者，酌予收錄，例如：「獲-穫」。

(二)「第 2a 層」、「第 2b 層」

1. 「第 2a 層」僅收錄「第 1 層」情形之外的大陸用簡體字；
2. 「第 2b 層」僅收錄第 1 及 2a 層之外的漢字，例如：日文漢字。

(三)「第 3 層」

收錄「第 1 層」、「第 2a 層」、「第 2b 層」情形之外的異體對應字。

四、對應表編排方式

(一) 初稿

初稿暫以表格方式呈現，並以交叉顏色區隔各組對應：

「Unicode 異體字對應表」初稿 (2009.7.1)				
(以不同顏色區隔各組對應)				
分級	字形	CCCII	Unicode	Big5
1	三	213024	04E09	A454
1	叁	393024	053C1	90E7
3	下	213025	04E0B	A455
3	𠂇	2D3025	04E05	8E47
3	上	213026	04E0A	A457
3	𠂈	2D3026	04E04	8E48

(二) 定稿

定稿將暫訂以【附件】樣式編排。

五、初稿意見徵詢

本對照表初稿擬於 7 月 1 日於 NBINet 網站公布後 3 個月內，徵詢圖書館及自動化系統廠商之修訂建議，請將意見傳送至國家圖書館書目資訊中心 nbinet@ncl.edu.tw，以作為初稿的修訂依據。初稿修訂完成後，擬於 10 月 10 日公布對應表定稿。

【附件】異體字對應表定稿樣式

層級 對應	第 1 層		第 2a 層			第 2b 層			第 3 層			
00441-CR				巖	严		巖					
00441-UN				056B4	04E25		053B3					
00441-CI				21375A	27375A		4B375A					
00533-CR										岸	岍	埠
00533-UN										05CB8	037C1	057BE
00533-CI										213B6C	393B6C	2D3B6C
00595-CR				廣	广		広			庑		
00595-UN				05EE3	05E7F		05E83			2220E		
00595-CI				213D2A	273D2A		333D2A			2D3D2A		
00684-CR				惱	恼		惱			惱	惱	
00684-UN				060F1	0607C		060A9			22649	226B0	
00684-CI				213E7E	273E7E		4B3E7E			2D3E7E	333E7E	
00740-CR	巔	巔		巔								
00740-UN	05DD4	05DD3		05DC5								
00740-CI	213C32	4B3C32		273C32								