

《佛說字經抄》舊版造字轉碼說明

中研院資訊所文獻處理實驗室
中央研究院語言所文獻語料小組
2009/05/21 丁致伶

轉碼主要工作是把檔案中的舊版造字轉換成 Windows XP 能支援的 Unicode 字形，Unicode 目前共收錄漢字 70194 個字，而 XP 只能支援 20902 個字（詳如表一），不支援之字將以構字式表達。例：造字編號 2189 的「𠄎」字，Unicode 編碼是 23C3D，由於 XP 並不支援，仍需使用構字式「毛△瞿」。

表一、Unicode 的字數及編碼區段

Unicode	新增字數	新增編碼區段	總字數	WinXP
1.1 版	20902	4E00-9FFF	20902	支援
3.0 版	6582	3400-4DFF	27484	不支援
3.1 版	42710	20000-2A6D6	70194	不支援

一、舊版造字轉碼分析：

《佛說字經抄》使用舊版造字 26 個，字頻 54 次，這 26 個造字中，23 個可轉成 Windows XP 能顯示的字，字頻 48 次；另外 3 個字必須轉成構字式，字頻 6 次。

轉碼完成製作轉碼分析表，請參考附件一《佛說字經抄》轉碼分析表，欄位說明如下：

- (一) 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
- (二) 造字：舊版造字。
- (三) 頻次：舊版造字在文件的出現次數。
- (四) Big5：造字的 Big5 碼。
- (五) Unicode：造字所對應的 Unicode 碼。

(六) WinXP：造字在 Windows XP 的對應字形。

(七) 構字式：Windows XP 無法對應字形改用構字式。

(八) 備註凡例：備註欄中記錄轉碼後字形及修改原因，凡例如下：

1. 異體字問題：為了使用者查詢和使用的方便，在處理異體字時最主要的方式是以標準字取代，除非是專有名詞或特殊情形，如：人名、地名等。例：造字編號 4768 的「妬」字，是「妒」的異體字，以標準字「妒」取代。
2. 錯字，以程式全部取代：與原書字形不符，並查詢教育部異體字字典確認非異體字後，皆歸類為錯字，以程式取代為正確字形。例：檔案中使用造字編號 4561 的「𠄎」字，而原書使用之字為「鹿」，查詢教育部異體字字典確認「𠄎」並非「鹿」之異體字，因此歸為錯字，以程式取代為正確字形「鹿」。
3. 待造字：Unicode 及漢字構形資料庫皆未收錄的舊漢籍造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 3893 的「𠄎」。

附件一、《佛說孛經抄》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
814	墓	5	8E5C	87C7	墓		
912	鷓	1	8EE0	9DC0	鷓		
972	犇	1	8F5D	7287	犇		
1053	傲	1	8FD0	6160	傲		
1420	嗶	2	9246	560A	嗶		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2044	踈	3	9642	8E08	踈		
2189	攮	1	96F5	23C3D		毛△瞿	
2190	瑜	1	96F6	6BFA	瑜		
2631	鴉	1	99D8	9D44	鴉		
2969	潛	1	9BF0	8B5B	潛		
3805	勅	4	8164	52C5	勅		
3814	却	2	816D	5374	却		
3819	廝	2	8172	53AE	廝		
3893		3	81DE			亻△妄	待造字。
4001	瘞	1	82AD	7602	瘞		
4150	訥	1	83A5	8A29	訥		
4184	迹	1	83C7	8FF9	迹		
4195	鉢	2	83D2	9262	鉢		
4203	隣	2	83DA	96A3	隣		
4515	呪	1	85D8	546A	呪		
4561		3	8647			𠃉〃厂	廐，錯字，以程式全部取

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
						𠂇比 \square	代。
4768	妬	4	8779	59AC	妬		妒，異體字問題。
4775	𡗗	3	87A2	59B7	𡗗		
5244	𡗗	2	8A7E	3738		薛 \triangle 女	
5300	鬪	5	8AD8	9B2A	鬪		
5702	厨	1	8D71	53A8	厨		

二、手動取代表說明：

《佛說字經抄》檔案中出現「●」字形，共計5字，比對原書後以Unicode字形或構字式手動取代，並製作「●」字手動取代表，請參考附件二。欄位說明如下：

1. 頁碼：「●」字形所在位置之原書頁碼。
2. 檔案原文摘錄：摘錄《佛說字經抄》檔案有「●」字形的文句。
3. 原書字形：「●」在原書中之字形，以此字取代檔案原文的「●」。
4. 備註：記錄修改相同詞句的次數或其他事項。

附件二、《佛說字經抄》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 730	𡗗●	毛 \triangle 翳	改為翳
p. 730	●·字欲就坐	毛 \triangle 登	

頁碼	檔案原文摘錄	原書字形	備註
p. 730	得訥●	酉 𠂔 永	
p. 732	●看垣牆	門 𠂔 視	改為闕
p. 733	酉● · 妒賢	酉 𠂔 永	