

《佛說寶網經》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
2009/05/07 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 1739 的「燿」字，Unicode 編碼是 3DFF，由於 XP 並不支援，仍需使用構字式「火𠂇僉」。

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：

《佛說寶網經》使用舊版造字 20 個，字頻 56 次，這 20 個造字中，19 個可轉成 Windows XP 能顯示的字，字頻 54 次；另外 1 個字必須轉成構字式，字頻 2 次。

轉碼完成製作轉碼分析表，請參考附件一《佛說寶網經》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 1978 的「讚」字，是「讚」的異體字，以標準字「讚」取代。

附件一、《佛說寶網經》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1536	麤	1	92DC	9E81	麤		
1739	燿	2	944B	3DFF		火𠩺僉	
1978	讚	7	95BF	8B83	讚		讚，異體字問題。
3814	却	3	816D	5374	却		
4001	瘞	3	82AD	7602	瘞		
4012	鼓	1	82B8	76B7	鼓		鼓，異體字問題。
4041	豎	2	82D5	7AEA	豎		
4090	脚	1	8347	811A	脚		
4146	覩	15	83A1	89A9	覩		睹，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4184	迹	1	83C7	8FF9	迹		
4195	鉢	1	83D2	9262	鉢		
4203	隣	1	83DA	96A3	隣		
4278	亘	1	8466	4E98	亘		
4515	呪	1	85D8	546A	呪		
4738	睺	2	875B	777A	睺		
5300	鬪	3	8AD8	9B2A	鬪		
5316	晉	1	8AE8	664B	晉		晉，異體字問題。
5437	并	6	8BC4	5E77	并		
5542	鑠	2	8C6E	93C1	鑠		
5555	迥	2	8C7B	9008	迥		

二、手動取代表說明：

《佛說寶網經》檔案中出現「●」字形，共計 35 字，比對原書後以 Unicod 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《佛說寶網經》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、《佛說寶網經》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 78	現光巨●數	姦	
p. 78	車●之胎	形王巨木	
p. 78	成就車●	形王巨木	
p. 78	億●無數	姦	
p. 79	億●百千	姦	
p. 79	千●諸天悉	姦	
p. 79	億百千●	姦	有 17 處
p. 80	●鬼人與非人	禘 厭	
p. 80	睹見億●	姦	
p. 81	尋見億●佛	姦	
p. 82	阿須倫健沓●	愬	
p. 83	百千億●	姦	有 2 處
p. 83	各十億●	姦	
p. 85	億百千兆●	姦	
p. 85	識億●宿	姦	
p. 86	無央數億●	姦	
p. 86	千●。而見	姦	
p. 86	阿●如來	閔	