

# 《七處三觀經》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  
2009/03/26 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 5206 的「顛」字，Unicode 編碼是 29573，由於 XP 並不支援，仍需使用構字式「寧𠄎頁」。

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：

《七處三觀經》使用舊版造字 10 個，字頻 62 次，這 10 個造字中，9 個可轉成 Windows XP 能顯示的字，字頻 61 次；另外 1 個字必須轉成構字式，字頻 1 次。

轉碼完成製作轉碼分析表，請參考附件一《七處三觀經》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 4134 的「衛」字，是「衛」的異體字，以標準字「衛」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. Unicode 字型呈現差異：Unicode 字型與舊版造字有些微差異，但只是字體風格差異，實際上仍為同一個字，因此仍取 Unicode 字型。如編號 3978 的「猪」字，Unicode 字型呈現為「猪」，實際上仍為同一字。

附件一、《七處三觀經》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
972	犇	2	8F5D	7287	犇		
3814	却	1	816D	5374	却		
3978	猪	2	8274	732A	猪		Unicode 字型呈現差異。
4134	衛	47	8373	885E	衛		衛，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4195	鉢	1	83D2	9262	鉢		
4256	鷄	2	8450	9DC4	鷄		雞，異體字問題。
4561		3	8647			𠃉 𠃊 𠃋 𠃌 𠃍 𠃎	龐，錯字，以程式全部取代。
5133	憇	1	89CE	6199	憇		
5206	顛	1	8A58	29573		寧 𠃏 頁	
5300	鬪	2	8AD8	9B2A	鬪		

## 二、手動取代表說明：

《七處三觀經》檔案中出現「●」字形，共計 7 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《七處三觀經》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、《七處三觀經》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
----	--------	------	----

頁碼	檔案原文摘錄	原書字形	備註
p. 877	紅汁●脹	臄	
p. 882	樹前芽栽●	扶	
p. 882	●拈便布	扶	
p. 882	不可●拈	扶	
p. 882	便●拈掘根	扶	
p. 882	為是●拈	扶	
p. 883	噫吐●下風	嚏	