

《興起行經》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
2009/04/16 丁玟伶

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 4856 的「𪗇」字，Unicode 編碼是 3CB2，由於 XP 並不支援，仍需使用構字式「疊𪗇毛」。

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：

《興起行經》使用舊版造字 29 個，字頻 105 次，這 29 個造字中，24 個可轉成 Windows XP 能顯示的字，字頻 91 次；另外 5 個字必須轉成構字式，字頻 14 次。

轉碼完成製作轉碼分析表，請參考附件一《興起行經》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。

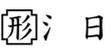
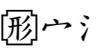
(八) 備註凡例：備註欄中記錄轉碼後字形及修改原因，凡例如下：

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 1513 的「粮」字，是「糧」的異體字，以標準字「糧」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 5500 的「滯」字，而原書使用之字為「恒」，「滯」為錯字，以程式取代為正確字形「恒」。

附件一、《興起行經》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
888	捺	7	8EC8	3B88		木 ㄩ 奈	
1513	粮	1	92C5	7CAE	粮		糧，異體字問題。
1536	麤	1	92DC	9E81	麤		
1926	覓	3	9569	8994	覓		
2172	攢	3	96E4	408E		矛 ㄩ 贊	
2524	饜	3	994B	994D	饜		
2994	揀	1	9C4A	6352	揀		
3805	勅	6	8164	52C5	勅		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3814	却	4	816D	5374	却		
3940	爵	2	824E	6B1D	爵		
4041	豎	1	82D5	7AEA	豎		
4085	耻	1	8342	803B	耻		
4089	脇	9	8346	8107	脇		
4090	脚	17	8347	811A	脚		
4139	袴	1	8378	88B4	袴		
4146	覩	4	83A1	89A9	覩		睹，異體字問題。
4195	鉢	6	83D2	9262	鉢		
4237	髡	9	83FC	9AE0	髡		
4287	織	1	846F	7E8E	織		
4515	呪	2	85D8	546A	呪		
4650	孳	2	86C2	4D38		麥  孚	
4768	妬	6	8779	59AC	妬		妒，異體字問題。
4775	妣	1	87A2	59B7	妣		
4856	毳	1	87F3	3CB2		疊  毛	
4889	毳	1	8855	3CAA		登  毛	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5105	埳	4	89B2	5860	埳		
5300	鬪	1	8AD8	9B2A	鬪		
5453	涅	6	8BD4	23D40		 	涅，異體字問題。
5500	滯	1	8C44	21A7B		 	恒，錯字，以程式全部取代。

二、手動取代表說明：

《興起行經》檔案中出現「●」字形，共計12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《興起行經》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、《興起行經》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 164	●破脚拇指	埳	
p. 165	縛驢●	馱	有3處
p. 165	無故見●	抔	

頁碼	檔案原文摘錄	原書字形	備註
p. 165	無故●困	𠂔	
p. 165	●困道士	𠂔	
p. 165	驢●此人	馱	
p. 165	即以驢●	馱	
p. 165	●困辟支佛	𠂔	
p. 172	毘毘●	翕𠂔毛	
p. 172	今得食●麥仁	扌𠂔鳥	