

《賢愚經》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
2009/04/20 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「𪛗」字，Unicode 編碼是 3CB2，由於 XP 並不支援，仍需使用構字式「疊𪛗毛」。

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：

《賢愚經》使用舊版造字 84 個，字頻 876 次，這 84 個造字中，77 個可轉成 Windows XP 能顯示的字，字頻 782 次；另外 7 個字必須轉成構字式，字頻 94 次。

轉碼完成製作轉碼分析表，請參考附件一《賢愚經》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」。
 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、《賢愚經》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
814	墓	4	8E5C	87C7	墓		
834	滄	1	8E70	712D	滄		
888	捺	41	8EC8	3B88		木𠂔奈	
1053	傲	3	8FD0	6160	傲		
1091	漚	4	8FF6	58CD	漚		
1111	瓚	1	904B	749D	瓚		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1237	稍	3	90EB	77DF	稍		
1400	窓	2	91F1	7A93	窓		
1513	糧	1	92C5	7CAE	糧		糧，異體字問題。
1533	剌	1	92D9	34DF		皮厶刂	
1926	覓	1	9569	8994	覓		
2042	䟽	2	9640	47FD		足厶流	
2044	踈	2	9642	8E08	踈		
2278	鋤	1	97B1	936E	鋤		
2408	鬧	1	9874	9599	鬧		鬧，異體字問題。
2506	飭	1	98F8	991D	飭		
2512	騶	1	98FE	26806		月厶巽	
2524	饜	17	994B	994D	饜		
2699	埠	2	9A5D	5813	埠		
3804	効	1	8163	52B9	効		
3805	勅	111	8164	52C5	勅		
3808	勾	24	8167	5303	勾		
3814	却	45	816D	5374	却		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3817	廁	3	8170	53A0	廁		廁，異體字問題。
3819	廐	1	8172	53AE	廐		廐，異體字問題。
3846	墻	2	81AF	58FB	墻		
3860	尠	4	81BD	5C20	尠		
3883	徧	3	81D4	5FA7	徧		
3884	忽	8	81D5	6031	忽		
3891	慙	3	81DC	6159	慙		
3900	挖	1	81E5	62D5	挖		
3940	欝	18	824E	6B1D	欝		
3944	氷	4	8252	6C37	氷		
3955	澗	3	825D	3D4E		澗	澗，異體字問題。
3966	烟	1	8268	70DF	烟		
3978	猪	1	8274	732A	猪		Unicode 字型呈現差異。
3980	獾	1	8276	7328	獾		
3984	瑠	2	827A	7460	瑠		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3992	嘗	1	82A4	751E	嘗		
4001	瘥	1	82AD	7602	瘥		
4016	着	1	82BC	7740	着		
4028	稟	13	82C8	7980	稟		
4040	竝	2	82D4	7ADD	竝		
4041	豎	20	82D5	7AEA	豎		
4081	翱	1	82FD	7FFA	翱		
4089	脇	4	8346	8107	脇		
4090	脚	20	8347	811A	脚		
4105	菓	5	8356	83D3	菓		
4134	衛	107	8373	885E	衛		衛，異體字問題。
4146	覩	66	83A1	89A9	覩		睹，異體字問題。
4160	適	2	83AF	8B81	適		
4173	躡	3	83BC	8EAD	躡		
4176	輒	5	83BF	8F19	輒		
4184	迹	10	83C7	8FF9	迹		
4195	鉢	47	83D2	9262	鉢		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4203	隣	4	83DA	96A3	隣		
4240	髻	2	8440	9AF4	髻		
4278	亘	1	8466	4E98	亘		
4342	倮	1	84C8	502E	倮		
4489		1	85BE			口会方	号，異體字問題。
4515	呪	22	85D8	546A	呪		
4716	塗	1	8745	57FF	塗		
4738	睺	14	875B	777A	睺		
4768	妬	19	8779	59AC	妬		妒，異體字問題。
4773	姉	8	877E	59C9	姉		Unicode 字型呈現差異。
4775	妣	1	87A2	59B7	妣		
4799	顛	1	87BA	9856	顛		
4856	毳	14	87F3	3CB2		疊  毛	
4889	毳	1	8855	3CAA		登  毛	
4904	尅	5	8864	5C05	尅		
4966	癩	1	88C4	764E	癩		癩，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4996	斲	12	88E2	65B5	斲		
5002	鞞	34	88E8	4A6D		鞞	
5300	鬪	5	8AD8	9B2A	鬪		
5305	晒	7	8ADD	661E	晒		
5447	躡	6	8BCE	8E83	躡		
5453	涅	58	8BD4	23D40		形: 日 工	涅，異體字問題。
5500	滌	5	8C44	21A7B		形: 彳 滌	恒，錯字，以程式全部取代。
5555	迥	1	8C7B	9008	迥		
5559	欸	6	8CA1	6B35	欸		
5702	厨	9	8D71	53A8	厨		
5706	敘	1	8D75	654D	敘		敘，異體字問題。
5742	澁	2	8DBB	6F81	澁		
5779	揿	1	8DE0	645A	揿		

二、手動取代表說明：

《賢愚經》檔案中出現「●」字形，共計66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《賢愚經》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、《賢愚經》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 349	涅槃·使此●	𠄎 𠄎朋	
p. 360	迦梨王●忍辱	𠄎	
p. 364	興兵●害民命	𠄎	
p. 364	往取●來	𠄎	
p. 369	取我●去	𠄎	
p. 372	白如玻●山	𠄎	
p. 372	王時眩●	目 𠄎 冒	
p. 372	金剛杵打玻●	𠄎	
p. 373	●●踊沒	足 𠄎 巨 足 𠄎 我	
p. 374	即自取藥·●	扌 𠄎 鳥	
p. 374	未●	扌 𠄎 鳥	
p. 375	煮以為●	食 𠄎 霍	

頁碼	檔案原文摘錄	原書字形	備註
p. 376	莫不厭●	𠂇 厶亥	
p. 377	波離阿●樓	少 厶兔	
p. 378	如●箭鈎	句 厶灬	待造字
p. 379	勢力·●尅	扌	
p. 379	拘執及●	翁 厶毛	
p. 379	身●痒故	形 疒 火 虫 口	待造字
p. 379	惡心刺●	扌	
p. 380	貪求●惜	恪	
p. 383	●者豈少	扌	
p. 385	●答婦言	聑	
p. 385	婦答●言	聑	
p. 385	婦語●	聑	
p. 385	值時●會	食 厶燕	待造字
p. 388	積為大●	艹 厶積	
p. 390	或復●裂	夨	
p. 390	●壞面者	夨	
p. 395	性慳●	恪	
p. 400	●馬二匹	馬 厶字	待造字
p. 401	●殺善人	扌	
p. 402	數倉●現穀	篇	
p. 403	倉●自滿	篇	

頁碼	檔案原文摘錄	原書字形	備註
p. 406	是我●	𠂔	
p. 409	儲●以待	𠂔亡大	待造字
p. 410	強弱相●	𠂔欠	
p. 410	●殺	𠂔	
p. 413	具●其事	𠂔	
p. 421	癡●拘癖	𠂔淺	待造字
p. 423	●破其面	𠂔	
p. 425	●師子	馬𠂔字	待造字
p. 426	一切皆送·●	口𠂔畢	
p. 428	●惱於我	𠂔心	
p. 428	不●民物	𠂔	
p. 429	爾●殺他兒	𠂔	
p. 429	●殺他	𠂔	
p. 429	虛妄·●名	𠂔	
p. 431	貪求慳●	𠂔	
p. 432	見其●甥兒	外𠂔男	待造字
p. 438	名號●迦羅	𠂔茶	
p. 438	●迦毘羅者	𠂔茶	
p. 440	見諸群●	利人水	
p. 442	優波●提	𠂔	有 8 處
p. 442	阿巴●提	𠂔	有 2 處

頁碼	檔案原文摘錄	原書字形	備註
p. 442	字為●提	𠄎	
p. 442	難陀●提	𠄎	
p. 442	優婆●	𠄎	