

《大寶積經密迹金剛力士會》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
2009/04/30 丁政伶

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：

《大寶積經密迹金剛力士會》使用舊版造字 38 個，字頻 369 次，這 38 個造字中，35 個可轉成 Windows XP 能顯示的字，字頻 365 次；另外 3 個字必須轉成構字式，字頻 4 次。

轉碼完成製作轉碼分析表，請參考附件一《大寶積經密迹金剛力士會》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. 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。
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「𠃉𠃊妄」。

附件一、《大寶積經密迹金剛力士會》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
972	犇	1	8F5D	7287	犇		
1400	窓	1	91F1	7A93	窓		
1536	麤	7	92DC	9E81	麤		
1978	讚	1	95BF	8B83	讚		讚，異體字問題。
2462	𠃉	1	98CC	4D2C		麥𠃉弋	
2502	颺	1	98F4	98C8	颺		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2524	饊	13	994B	994D	饊		
2885	叫	1	9B7A	544C	叫		
3805	勅	2	8164	52C5	勅		
3814	却	6	816D	5374	却		
3860	尠	1	81BD	5C20	尠		
3893		2	81DE			亻 ㄩ 妄	待造字。
3966	烟	1	8268	70DF	烟		
3978	猪	1	8274	732A	猪		Unicode 字型 呈現差異。
3980	獍	1	8276	7328	獍		
4001	瘞	2	82AD	7602	瘞		
4002	瘡	1	82AE	7607	瘡		
4004	瘠	1	82B0	763B	瘠		
4041	豎	4	82D5	7AEA	豎		
4090	脚	2	8347	811A	脚		
4134	衛	1	8373	885E	衛		衛，異體字問 題。
4146	覩	39	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	170	83C7	8FF9	迹		
4195	鉢	10	83D2	9262	鉢		
4203	隣	2	83DA	96A3	隣		
4254	鷓	1	844E	9D7E	鷓		
4267	鼈	1	845B	9F08	鼈		
4495	麩	1	85C4	9EA8	麩		Unicode 字型 呈現差異。
4515	呪	50	85D8	546A	呪		
4546	鴟	1	85F7	9D76	鴟		
4716	塗	1	8745	57FF	塗		
4768	妬	2	8779	59AC	妬		妒，異體字問 題。
4856	疊	1	87F3	3CB2		疊  毛	
5133	憇	10	89CE	6199	憇		
5300	鬪	7	8AD8	9B2A	鬪		
5316	晉	9	8AE8	664B	晉		晉，異體字問 題。
5437	并	11	8BC4	5E77	并		
5472	瓦	2	8BE7		瓦		

二、手動取代表說明：

《大寶積經密迹金剛力士會》檔案中出現「●」字形，共計38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《大寶積經密迹金剛力士會》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、《大寶積經密迹金剛力士會》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 42	●。從無限	𠄎	
p. 42	●施。執像蜜天	龍	
p. 42	捷沓●	𠄎	
p. 44	眾生多所●業	恪	
p. 47	有八千●	𠄎	
p. 50	億載兆●人	𠄎	
p. 51	七十六●人	𠄎	
p. 51	作七寶●顯現	𠄎𠄎瓦	
p. 52	舉瓶著●上	𠄎𠄎瓦	
p. 54	三萬兆●	𠄎	
p. 56	紺色從天中●	立𠄎寺	
p. 56	膝●無瞋休息	𠄎	

頁碼	檔案原文摘錄	原書字形	備註
p. 57	億百千●	𡗗	有 6 處
p. 58	木密●彼亦	𡗗	
p. 58	亦●印氏	𡗗	
p. 61	化六十●人	𡗗	
p. 62	名香木●	𡗗	有 2 處
p. 64	數百千億●	𡗗	
p. 64	應時八十●魔	𡗗	
p. 64	十億●鬼神	𡗗	
p. 65	六十八萬億●	𡗗	
p. 68	生阿●佛土	𡗗	
p. 68	其阿●佛	𡗗	
p. 68	水精車●馬	𡗗王巨木	
p. 70	碑●為地	𡗗王巨木	
p. 71	棄擔●心	揆	
p. 72	博聞無所愛●	恪	
p. 74	壽命不以為●	恪	
p. 75	●如須彌山	木𡗗園	
p. 77	從阿●佛	𡗗	
p. 78	不以●惜	恪	
p. 80	車●馬瑙	𡗗王巨木	