

《正法念處經》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
2009/05/19 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：

《正法念處經》使用舊版造字 154 個，字頻 6981 次，這 154 個造字中，122 個可轉成 Windows XP 能顯示的字，字頻 6507 次；另外 32 個字必須轉成有構字符號的構字式，字頻 474 次。

轉碼完成製作轉碼分析表，請參考附件一《正法念處經》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〃厂曲比[◻]」字，而原書使用之字為「鹿」，查詢教育部異體字字典確認「𠄎〃厂曲比[◻]」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。
 4. **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 4547 的「口𠄎美」。
 5. **構字式部件，轉為 Big5 編碼字形**：該字在檔案中是為構字式部件，雖有 XP 支援的 Unicode 字形，但因構字式不支援 Unicode 部件，仍需轉為 Big5 編碼的字形。例：編號 379 的「支」字，雖有 XP 支援的 Unicode 字形「支」字，但因該字在檔案中為構字式部件，構字式不支援 Unicode 部件，仍需轉為 Big5 編碼的「支」字。

附件一、《正法念處經》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
296	厶	29	FBEC			厶	
297	厶	170	FBED			厶	
298	厶	32	FBEE			厶	
299	形	60	FBEF			形	
300	口	60	FBF0			口	
379	支	1	FCA2	6534		支	構字式部件， 轉為Big5編碼 字形。
406	苙	4	FCBD			苙	
658	疒	32	FE5D	7592		疒	構字式部件， 轉為Big5編碼 字形。
660	丂	1	FE5F	4E02		丂	
716	夂	1	FEB9	7676		夂	構字式部件， 轉為Big5編碼 字形。
717	足	4	FEBA			足	
746	丩	15	FED7	5FC4		丩	構字式部件， 轉為Big5編碼

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							字形。
752	𠂇	6	FEDD			𠂇	
768	𠂈	8	FEEF	8279		𠂈	構字式部件，轉為Big5編碼字形。
770	扌	11	FEEF	624C		扌	構字式部件，轉為Big5編碼字形。
772	辶	1	FEF1	8FB6		辶	構字式部件，轉為Big5編碼字形。
777	宀	5	FEF6	5B80		宀	構字式部件，轉為Big5編碼字形。
814	墓	6	8E5C	87C7	墓		
1016	揸	2	8FAB	22D24		扌 𠂇 棠	
1053	傲	4	8FD0	6160	傲		
1219		1	90D9			宀 𠂇 巢	窠，錯字，以程式全部取代。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1237	稍	11	90EB	77DF	稍		
1383	糠	1	91E0	7A45	糠		
1387	瞽	2	91E4	7796	瞽		
1400	窓	5	91F1	7A93	窓		
1513	糧	1	92C5	7CAE	糧		糧，異體字問題。
1533	剌	1	92D9	34DF		皮厶刂	
1536	麇	61	92DC	9E81	麇		
1772	藜	1	946C	4527		艹厶梨	
1864	蛄	1	94EA	872B	蛄		
1937	詢	1	9574	8A3D	詢		
1951	味	1	95A4	20C0C		口厶末	味，異體字問題。
1978	讚	192	95BF	8B83	讚		讚，異體字問題。
1999	贊	21	95D4	8CDB	贊		贊，異體字問題。
2044	踈	2	9642	8E08	踈		
2147	蟞	1	96CB	87EE	蟞		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2162	醜	17	96DA	918E	醜		
2298	殼	3	97C5	3C89		殼△卵	
2325	潤	9	97E0	77A4	潤		
2384		1	985C			几△黑	利△黑，錯字，以程式全部取代。
2408	鬧	8	9874	9599	鬧		鬧，異體字問題。
2439	磁	1	98B5	7920		石△慈	
2448	隼	2	98BE	96BD	隼		
2524	饜	5	994B	994D	饜		
2539	脍	3	995A	43F6		月△望	
2616	鱈	1	99C9	9C6F	鱈		
2631	鴉	17	99D8	9D44	鴉		
2651	鸚	3	99EC	9E1C	鸚		
2699	埠	2	9A5D	5813	埠		
2719	函	3	9A71	51FE	函		
2739	忙	1	9AA7	607E	忙		
2835	梯	1	9B48	632E	梯		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2885	叫	1	9B7A	544C	叫		
2975	迓	1	9BF6	8FEF	迓		
3143	潰	1	9D42	6FFD	潰		
3202	癩	3	9D7D	7636	癩		
3255	善	447	9DD4	5584	善		
3792	涼	1	8157	51C9	涼		涼，異體字問題。
3800		2	815F			杀  	刹，異體字問題。
3805	勅	29	8164	52C5	勅		
3808	勾	1	8167	5303	勾		
3814	却	17	816D	5374	却		
3817	廁	11	8170	53A0	廁		廁，異體字問題。
3824	叙	2	8177	53D9	叙		
3829	咏	40	817C	548F	咏		
3858	冤	1	81BB	5BC3	冤		冤，異體字問題。
3860	尠	2	81BD	5C20	尠		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3864	峯	319	81C1	5CEF	峯		
3931	椀	1	8245	6900	椀		
3940	欝	127	824E	6B1D	欝		
3944	氷	10	8252	6C37	氷		
3955	澗	13	825D	3D4E		澗	澗，異體字問題。
3966	烟	17	8268	70DF	烟		
3975	牕	1	8271	7255	牕		
3978	猪	23	8274	732A	猪		Unicode 字型呈現差異。
3979	猫	4	8275	732B	猫		
3980	猿	1	8276	7328	猿		
3984	瑠	2	827A	7460	瑠		
3989	璿	4	82A1	74C8	璿		
3992	嘗	5	82A4	751E	嘗		
4001	瘞	4	82AD	7602	瘞		
4004	瘵	1	82B0	763B	瘵		
4010	臯	1	82B6	7690	臯		
4040	竝	9	82D4	7ADD	竝		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4041	豎	19	82D5	7AEA	豎		
4086	擲	1	8343	22D8F		扌 𠂔 剝	
4089	脇	27	8346	8107	脇		
4090	脚	37	8347	811A	脚		
4105	菓	7	8356	83D3	菓		
4124	蠱	1	8369	8771	蠱		
4132		3507	8371			血 𠂔 𠂔	眾，異體字問題。
4134	衛	1	8373	885E	衛		衛，異體字問題。
4146	覩	15	83A1	89A9	覩		睹，異體字問題。
4156	諂	1	83AB	8B1F	諂		
4184	迹	2	83C7	8FF9	迹		
4185	邊	1	83C8	9037	邊		
4193	鈎	4	83D0	920E	鈎		鈎，異體字問題。
4195	鉢	216	83D2	9262	鉢		
4203	隣	6	83DA	96A3	隣		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4213	鞞	1	83E4	97BE	鞞		
4226	凜	2	83F1	98E1	凜		
4240	鬚	2	8440	9AF4	鬚		
4256	鷄	16	8450	9DC4	鷄		雞，異體字問題。
4260	麇	7	8454	9E9E	麇		
4267	鼈	5	845B	9F08	鼈		
4271	齧	1	845F	9F69	齧		
4308	刼	1	84A6	5227	刼		
4325	賒	2	84B7	8CD6	賒		
4356	鞞	1	84D6	976B	鞞		
4432		5	8563			形糸厂 墨	纏，異體字問題。
4448		26	8573			形糸厂 墨	纏，異體字問題。
4495	麪	6	85C4	9EA8	麪		Unicode 字型呈現差異。
4496	段	1	85C5	53DA	段		
4513	眇	1	85D6	7717	眇		眇，錯字，以

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							程式全部取代。
4515	呪	22	85D8	546A	呪		
4528	嶼	5	85E5	5D60	嶼		
4531	攢	6	85E8	6505	攢		
4547		1	85F8			口厶美	待造字。
4551	嶠	2	85FC	21F0F		山厶間	
4561		10	8647			𠂇 厶 曲比 𠂇	麤，錯字，以程式全部取代。
4716	塗	1	8745	57FF	塗		
4738	睺	135	875B	777A	睺		
4739	鞞	5	875C	97DB	鞞		
4742		1	875F			衣厶執	褻，異體字問題。
4768	妬	82	8779	59AC	妬		妒，異體字問題。
4778	鞞	1	87A5	97B4	鞞		
4803	腭	3	87BE	816D	腭		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4856	氈	10	87F3	3CB2		疊厶毛	
4904	尅	1	8864	5C05	尅		
4947	毳	2	88B1	2917E		雲厶甚	
4982	囊	2	88D4	3C06		𠂇厶隹	
4996	斲	3	88E2	65B5	斲		
5031	闋	1	8946	95B4	闋		
5074	𠂇	1	8971	5C2B	𠂇		
5075	𠂇	2	8972	72AD		𠂇	構字式部件， 轉為Big5編碼 字形。
5108	愾	6	89B5	60FD	愾		
5133	憇	82	89CE	6199	憇		
5192	隙	1	8A4A	969F	隙		
5195		14	8A4D			𠂇木畝 且	檀，異體字問 題。
5246		1	8AA2			𠂇正人 月	踰，異體字問 題。
5300	鬪	210	8AD8	9B2A	鬪		
5437	并	46	8BC4	5E77	并		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5439	栢	3	8BC6	6822	栢		
5453	涅	322	8BD4	23D40		形 日 工 口	涅，異體字問題。
5496	羸	80	8C40	7F82	羸		
5516		10	8C54			〇 束	棘，異體字問題。
5644	浹	2	8CF6	3CEC		氵 厶 疋	
5666	汚	5	8D4D	6C5A	汚		汚，異體字問題。
5707	棧	8	8D76	6830	棧		
5709	蟀	5	8D78	872F	蟀		
5714	瓮	1	8D7D	74EB	瓮		
5724	穡	1	8DA9	7A38	穡		
5742	澁	26	8DBB	6F81	澁		
5744	萑	7	8DBD	84A6	萑		Unicode 字型呈現差異。

二、手動取代表說明：

《正法念處經》檔案中出現「●」字形，共計1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《正法念處經》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、《正法念處經》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
P. 223	猶如●氷者	鑽	