

《太平經合校》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
2008/12/03 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 4855 的「汩」字，Unicode 編碼是 3CC5，由於 XP 並不支援，仍需使用構字式「汩 𠄎不」。

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：

《太平經合校》使用舊版造字 60 個，字頻 991 次，這 60 個造字中，56 個可轉成 Windows XP 能顯示的字，字頻 978 次；另外 4 個字必須轉成構字式，字頻 13 次。

轉碼完成製作轉碼分析表，請參考附件一《太平經合校》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 982 的「胃」字，是「胸」的異體字，以標準字「胸」取代。

附件一、《太平經合校》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
982	胃	12	8F67	80F7	胃		胸，異體字問題。
1023	調	1	8FB2	8AB7	調		
1044	儻	1	8FC7	511B	儻		
1513	糧	2	92C5	7CAE	糧		糧，異體字問題。
1664	腸	1	93BF	8193	腸		
2500	涑	18	98F2	6E18	涑		
2699	埠	1	9A5D	5813	埠		
3787	冢	3	8152	51A2	冢		
3799	刼	1	815E	5226	刼		
3804	効	32	8163	52B9	効		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3805	勅	2	8164	52C5	勅		
3814	却	33	816D	5374	却		
3836	喻	1	81A5	564F	喻		
3846	壻	4	81AF	58FB	壻		
3850	姪	19	81B3	59D9	姪		
3858	寃	5	81BB	5BC3	寃		寃，異體字問題。
3869	币	3	81C6	5E00	币		
3876	廻	1	81CD	5EFB	廻		
3883	徧	15	81D4	5FA7	徧		
3891	慙	7	81DC	6159	慙		
3963	炁	39	8265	7081	炁		
3974	牀	6	8270	7240	牀		
3996	疎	27	82A8	758E	疎		
4038	竈	2	82D2	7AC8	竈		
4041	豎	1	82D5	7AEA	豎		
4055	粃	6	82E3	7C83	粃		
4080	羣	54	82FC	7FA3	羣		
4089	脇	4	8346	8107	脇		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4132		255	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	1	8373	885E	衛		衛，異體字問題。
4146	覩	117	83A1	89A9	覩		睹，異體字問題。
4160	適	8	83AF	8B81	適		
4184	迹	1	83C7	8FF9	迹		
4203	隣	15	83DA	96A3	隣		鄰，異體字問題。
4240	髻	1	8440	9AF4	髻		
4242	鬪	55	8442	9B2D	鬪		
4267	鼈	1	845B	9F08	鼈		
4307	罵	1	84A5	99E1	罵		罵，異體字問題。
4308	劫	1	84A6	5227	劫		
4342	倮	1	84C8	502E	倮		
4460	勅	160	85A1	52D1	勅		
4469	勺	9	85AA	5304	勺		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4515	呪	3	85D8	546A	呪		
4768	妬	5	8779	59AC	妬		妒，異體字問題。
4814	健	1	87C9	5FA4	健		
4855	汩	9	87F2	3CC5		汩 ㄩ不	
4870	壘	5	8842	97F2	壘		
4904	尅	7	8864	5C05	尅		
4987	躡	1	88D9	4831		足 ㄩ屬	
5126	暮	1	89C7	8B29	暮		
5169	危	6	89F2	6239	危		
5281	餓	1	8AC5	994A	餓		
5302	昏	1	8ADA	662C	昏		
5389	竿	1	8B72	7B07	竿		
5395	規	1	8B78	4093		矢 ㄩ見	
5439	栢	1	8BC6	6822	栢		
5459	缺	2	8BDA	21647		垂 ㄩ夬	
5706	敘	2	8D75	654D	敘		敘，異體字問題。
5724	穉	2	8DA9	7A38	穉		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5860	无	15	C6D3	65E0	无		

二、手動取代表說明：

《太平經合校》檔案中出現「●」字形，共計78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《太平經合校》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、《太平經合校》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 35	心大悲而恐●	忄 厶 亥	
p. 52	可●哉	忄 厶 亥	有 28 處
p. 116	可不●哉	忄 厶 亥	
p. 187	恐●驚	忄 厶 亥	
p. 189	且●且喜	忄 厶 亥	
p. 234	天師●哀憐愚生	天 厶 并	
p. 244	唯真人迺知一●	忄 厶 亥	
p. 246	真人知蚤●	忄 厶 亥	
p. 248	但見真人●●	倝 倝	

頁碼	檔案原文摘錄	原書字形	備註
p. 249	吾甚●哉	忄 厶 亥	
p. 249	子知●	忄 厶 亥	有 2 處
p. 252	吾甚驚●心痛	忄 厶 亥	
p. 266	國家諸侯王之●也	士 厶 聃	
p. 266	帝王女弟之●也	士 厶 聃	
p. 268	轉而異●	辭	原字不清，以註釋修改。
p. 276	●●無慮	倝 倝	
p. 297	故吾文□□●●	倝 倝	
p. 298	何及於●●當前乎哉	倝 倝	
p. 306	令賢聖惶●	忄 厶 亥	
p. 317	俱食氣迺●不通	口 厶 節	
p. 317	而氣●不通者	口 厶 節	
p. 357	愚生甚●	忄 厶 亥	
p. 360	子知●畏天談	忄 厶 亥	
p. 397	共辨難何●●	恂 恂	
p. 397	雖●●	恂 恂	
p. 397	天下●●	恂 恂	
p. 406	使天下●●多變諍	恂 恂	
p. 535	復益悵然有慚●	忄 厶 亥	
p. 622	善●，恣其所得	睪	

頁碼	檔案原文摘錄	原書字形	備註
p. 640	子知早●	𠂇 𠂇亥	
p. 640	不知●	𠂇 𠂇亥	
p. 661	真人知●	𠂇 𠂇亥	
p. 661	子不●	𠂇 𠂇亥	
p. 663	子知驚●	𠂇 𠂇亥	
p. 663	不知*驚●	𠂇 𠂇亥	
p. 677	甚惶甚●	𠂇 𠂇亥	
p. 677	子知惶且●	𠂇 𠂇亥	
p. 720	反●肝膽為發怒	𠂇 𠂇亥	
p. 725	廣大可●	𠂇 𠂇亥	
p. 737	揮魁●之劍	鬼△勺	
p. 737	鱗絡金線●	糸 𠂇區	
p. 742	羣神恐●	𠂇 𠂇亥	