

《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

中研院史語所漢籍電子文獻工作小組

2008/2/22 陳建安 製作

一、《太平經合校》一書(太平經合校.xml)使用舊版造字 73 個，字頻 1053 次，詳如附件一。這 73 個造字中，67 個可轉成 Windows XP 能顯示的字，字頻 784 次；另外 6 個字必須轉成構字式，字頻 269 次。

二、附件一的造字分析表說明如下：

甲、編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、造字：舊版造字

丙、字頻(txt)：造字在「.txt」文件的出現次數

丁、字頻(xml)：造字在「.xml」文件的出現次數

戊、Big5：造字的 Big5 碼

己、Unicode：造字所對應的 Unicode 碼

庚、WinXP：造字在 Windows XP 的對應字形

辛、備註凡例：

- 1、校對問題，舊版漢籍錯字，可用程式全部取代：在舊版漢籍電子文獻中即存在的錯字，因校對時的疏漏而未更正，持續留存在新版漢籍電子文獻中；若該造字的所有頻次，皆屬於錯誤使用的錯字情形，可以用程式全部取代為正確字形。如編號 3919 的「𠂔𠂔毛」字，原字為「无」。
- 2、異體字問題：新版漢籍考量到使用者檢索及使用時的便利性，將用字原則改為除專詞等特殊情形之外，一律改用標準字呈現。如編號 4132 的「血𠂔𠂔」係「眾」字之異體，故以「眾」字取代。又編號 1513 的「糧」係「糧」字之簡體，亦以正體之「糧」字取代。
- 3、Unicode 字型呈現差異：Unicode 字型與舊漢籍造字有些微差異，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 4158 的「譌」字，Unicode 字型呈現為「譌」，實際上仍為同一字。

三、Unicode 目前收錄的漢字總數為 70194，分屬於三個不同區段，詳如表一。Windows XP 只支援 CJK 認同表意文字區的 20902 個字，內碼為 4E00-9FFF。所以如造字編號 4855 的「汩」字，Unicode

碼為 3CC5，Windows XP 並不支援，故仍須使用構字式「𠂇 𠂇不」。

表一、Unicode 的字數及編碼區段

Unicode	字集子集合	新增字數	新增編碼區段	總字數	WinXP
1.1 版	CJK 認同表意文字區	20902	4E00-9FFF	20902	支援
3.0 版	CJK 認同表意文字擴充 A 區	6582	3400-4DFF	27484	不支援
3.1 版	CJK 認同表意文字擴充 B 區	42710	20000-2A6D6	70194	不支援

四、《太平經合校》中未確定字形「●」共 82 處，詳細處理內容可參考附件二的未確定字形「●」手動取代表。附件二《太平經合校》未確定字形「●」手動取代表欄位說明如下：

- 甲、檔案原文摘錄：摘錄檔案(太平經合校.xml)中包含未確定字形「●」的文句段落，其中紅字底線的部分，為需要手動取代的字形。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。
- 乙、手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。
- 丙、備註：註明其他備註事項或相同文句在檔案中出現的次數，未註明者表示只出現一次。

附件一、《太平經合校》造字分析表

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
982	𠂇	12	12	8F67	80F7	𠂇		
1014	玠	1	1	8FA9	73CE	玠		
1023	調	1	1	8FB2	8AB7	調		
1044	儻	1	1	8FC7	511B	儻		
1187	盖	1	1	90B9	76D6	盖		
1238	𠂇	1	1	90EC	77E6	𠂇		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
1513	糧	2	2	92C5	7CAE	糧		糧，異體字問題。
1664	腸	1	1	93BF	8193	腸		
2500	滌	18	18	98F2	6E18	滌		
2699	埤	1	1	9A5D	5813	埤		
3787	豕	3	3	8152	51A2	豕		
3799	刂	1	1	815E	5226	刂		
3804	効	35	35	8163	52B9	効		
3805	勅	2	2	8164	52C5	勅		
3814	却	40	40	816D	5374	却		
3836	喻	1	1	81A5	564F	喻		
3846	壻	4	4	81AF	58FB	壻		Unicode 呈現差異。
3850	姪	19	19	81B3	59D9	姪		
3858	寃	5	5	81BB	5BC3	寃		
3864	峯	1	1	81C1	5CEF	峯		
3869	帀	3	3	81C6	5E00	帀		
3876	廻	1	1	81CD	5EFB	廻		
3883	徧	16	16	81D4	5FA7	徧		
3891	慙	7	7	81DC	6159	慙		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
3919	𧈧	1	1	81F8	23C0B		𧈧𠂇毛	无，校對問題，舊版漢籍錯字，可用程式全部取代。
3961	灾	1	1	8263	707E	灾		
3963	𧈧	45	45	8265	7081	𧈧		
3974	牀	6	6	8270	7240	牀		
3996	疎	27	27	82A8	758E	疎		
4016	着	5	5	82BC	7740	着		著，異體字問題。
4038	竈	2	2	82D2	7AC8	竈		
4041	豎	1	1	82D5	7AEA	豎		
4055	𧈧	6	6	82E3	7C83	𧈧		
4080	羣	58	58	82FC	7FA3	羣		Unicode 呈現差異。
4089	脇	4	4	8346	8107	脇		
4104	莅	1	1	8355	8385	莅		
4132		255	255	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	1	1	8373	885E	衛		
4146	覩	118	118	83A1	89A9	覩		
4158	譌	11	11	83AD	8B4C	譌		Unicode 呈現差異。
4160	適	8	8	83AF	8B81	適		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4184	迹	3	3	83C7	8FF9	迹		
4190	邳	2	2	83CD	90C4	邳		
4203	隣	15	15	83DA	96A3	隣		
4240	髯	1	1	8440	9AF4	髯		
4242	鬪	55	55	8442	9B2D	鬪		Unicode 呈現差異。
4267	鼈	1	1	845B	9F08	鼈		
4307	罵	1	1	84A5	99E1	罵		
4308	刼	1	1	84A6	5227	刼		
4342	倮	1	1	84C8	502E	倮		
4460	勅	163	163	85A1	52D1	勅		
4469	勻	9	9	85AA	5304	勻		
4507	吳	1	1	85D0	5434	吳		吳，異體字問題。
4515	呪	3	3	85D8	546A	呪		
4768	妬	5	5	8779	59AC	妬		
4814	健	1	1	87C9	5FA4	健		
4855	汩	9	9	87F2	3CC5		汩 𣶒 不	
4870	壑	5	5	8842	97F2	壑		
4904	尅	7	7	8864	5C05	尅		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4987	躡	1	1	88D9	4831		足厶屬	
5126	蓐	1	1	89C7	8B29	蓐		
5169	𠂔	6	6	89F2	6239	𠂔		
5281	餽	1	1	8AC5	994A	餽		
5302	昏	1	1	8ADA	662C	昏		
5389	竿	1	1	8B72	7B07	竿		
5395	規	1	1	8B78	4093		矢厶見	規，異體字問題。
5439	栢	1	1	8BC6	6822	栢		
5459	𠂔	2	2	8BDA	21647		垂厶夊	
5706	敍	7	7	8D75	654D	敍		
5724	穡	2	2	8DA9	7A38	穡		
5752	摠	1	1	8DC5	63D4	摠		
5753	潛	1	1	8DC6	6F5C	潛		潛，異體字問題。
5860	无	17	17	C6D3	65E0	无		

附件二、《太平經合校》未確定字形「●」手動取代表

檔案原文摘錄	手動取代結果	備註
殺●	殺杀厶支	
心大悲而恐●	心大悲而恐忄厶亥	
可●哉，可●哉（全部取代）	可忄厶亥哉，可忄厶亥哉（全部取代）	2
可不●哉	可不忄厶亥哉	

檔案原文摘錄	手動取代結果	備註
可●哉！（全部取代）	可忄△亥哉！（全部取代）	24
恐●驚	恐忄△亥驚	
且●且喜	且忄△亥且喜	
天師●哀憐愚生	天師天△并哀憐愚生	
唯真人迺知一●	唯真人迺知一忄△亥	
真人知蚤●	真人知蚤忄△亥	
但見真人●●	但見真人倻倻	
吾甚●哉	吾甚忄△亥哉	
子知●	子知忄△亥	
吾甚驚●心痛	吾甚驚忄△亥心痛	
是國家諸侯王之●	是國家諸侯王之士△耳	
是帝王女弟之●	是帝王女弟之士△耳	
轉而異●	轉而異辭	（註）
●●無慮	倻倻無慮	
故吾文□□●●	故吾文□□倻倻	
何及於●●當前乎哉	何及於倻倻當前乎哉	
令賢聖惶●	令賢聖惶忄△亥	
俱食氣迺●不通	俱食氣迺口△節不通	
而氣●不通者	而氣口△節不通者	
●。」「子知●	忄△亥。」「子知忄△亥	
子知●畏天談	子知忄△亥畏天談	
共辨難何●●	共辨難何恂恂	
雖●●，無益也	雖恂恂，無益也	
天下●●（全部取代）	天下恂恂（全部取代）	2
復益悵然有慚●	復益悵然有慚忄△亥	
善●，恣其所得	善腴，恣其所得	
子知早●	子知早忄△亥	
不知●，死之根	不知忄△亥，死之根	
真人知●	真人知忄△亥	
子不●	子不忄△亥	
子知驚●	子知驚忄△亥	
不知＊驚●	不知＊驚忄△亥	
甚惶甚●	甚惶甚忄△亥	
子知惶且●	子知惶且忄△亥	
反●肝膽為發怒	反忄△亥肝膽為發怒	
山川之廣大可●	山川之廣大可忄△亥	

檔案原文摘錄	手動取代結果	備註
揮魁●之劍	揮魁鬼△勺之劍	
鱗絡金線●	鱗絡金線糸△區	
羣神恐●	羣神恐忄△亥	
解天●九人決	解天宀△西△電九人決	
善言教人增●決	善言教人增筴決	
象文行增●決	象文行增筴決	

(註) 原字殘缺不全，註解謂似「辭」字，故暫時以「辭」字取代。