《前漢紀》舊版造字轉碼說明

中研院資訊所文獻處理實驗室 96.7.31 陳建安製作

一、 《前漢紀》(前漢紀.xml)使用舊版造字 113 個，字頻 977 次，詳如附件一。這 113 個造字中，103 個可轉成 Windows XP 可顯示的字，字頻 699 次；另外 20 個字必須轉成構字式，字頻 278 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：造字在文件中的出現次數
丁、 Big5：造字的 Big5 碼
戊、 Unicode：造字所對應的 Unicode 碼
己、 WinXP：造字在 Windows XP 的對應字形
庚、 構字式： Windows XP 若無對應字形，則改採用構字式
辛、 備註凡例：

1、部分錯字：部分字形與原書字形不符，用正確的字取代原造字，如編號 2490 的「頟」字，書中共出現 2 次，但原書字形有 1 次為「頟」字，是部分錯字，必須手動取代。

2、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 4773 的「姊」字，Unicode 字型呈現為「姊」，實際上仍為同一字。

3、待造字：舊版漢籍造字字形為電腦缺字，且漢字構形資料庫尚未收錄此字形，等待補造字中，所以「造字」欄空白無法看到字形。如編號 2811 的「⿱丣⿱亼田」。

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。目前 Windows XP 只支援 CJK 認同表意文字區的 20902 個字，內碼為 4E00-9FFF。所以造字編號 4825 的「⿵門⿱亼⿰⿱亼田」字，Unicode 編碼為 28D66，由於 Windows XP 並不支援，仍須改用構字式「門⿱亼⿰⿱亼田」。

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、 造字分析表「備註」欄標示「手動取代」者，詳細取代內容可參考附件二的手動取代表。附件二的手動取代表說明如下：

甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
乙、 造字：舊漢籍造字字形
丙、 部件：構字式字形，通常當部件使用。
丁、 檔案原文摘錄：摘錄檔案(前漢紀.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。
戊、 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。

五、 《前漢紀》(前漢紀.xml)中部分字形顯示為「●」，已對照原書字形更正，詳細更正內容可參考附件三的「●」字形手動取代表。附件三「●」字形手動取代表的說明同前丁、戊兩項。

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
832	沲	1	1	8E6E	6CB2	沲		
852	狢	1	1	8EA4	72E2	狢		
957	遫	1	1	8F4E	906C	遫		
1024	虵	3	3	8FB3	8675	虵		
1039	稟	1	1	8FC2	69C0	稟		
1067	棊	2	2	8FDE	68CA	棊		
1077	稟	1	1	8FE8	7A3E	稟		
1079	總	1	1	8FEA	7DEB	總		Unicode 字型呈現差異
1565	緜	4	4	92F9	7DDC	緜		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
1591	羌	4	4	9354	7F97	羌		
1885	蠭	2	2	9540	882D	蠭		
1906	褎	12	12	9555	88E6	褎		
2424	桀	1	1	98A6	6909	桀		
2490	頟	2	2	98E8			名⿰⿱八人頁	頟部分錯字，手動取代
2651	鸜	2	2	99EC	9E1C	鸜		
2811		23	23	9AEF			帀⿱亼田	待造字
3787	冢	27	27	8152	51A2	冢		
3789	冲	1	1	8154	51B2	冲		
3792	凉	1	1	8157	51C9	凉		
3804	効	3	3	8163	52B9	効		
3805	勅	2	2	8164	52C5	勅		
3813	卬	8	8	816C	536D	卬		
3814	却	3	3	816D	5374	却		
3819	厮	3	3	8172	53AE	厮		
3829	咏	1	1	817C	548F	咏		
3830	咤	2	2	817D	54A4	咤		
3840	坂	6	6	81A9	5742	坂		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
3846	壻	4	4	81AF	58FB	壻		
3850	姙	1	1	81B3	59D9	姙		
3858	寃	6	6	81BB	5BC3	寃		
3869	帀	1	1	81C6	5E00	帀		
3883	徧	6	6	81D4	5FA7	徧		
3891	慙	3	3	81DC	6159	慙		
3930	椁	7	7	8244	6901	椁		
3931	椀	1	1	8245	6900	椀		
3949	涖	2	2	8257	6D96	涖		
3974	牀	8	8	8270	7240	牀		
3976	犂	4	4	8272	7282	犂		
3984	瑠	1	1	827A	7460	瑠		
3996	疎	16	16	82A8	758E	疎		
4015	眦	3	3	82BB	7726	眦		
4034	穉	3	3	82CE	7A49	穉		
4038	竈	9	9	82D2	7AC8	竈		
4040	竝	65	65	82D4	7ADD	竝		
4070	緤	2	2	82F2	7DE4	緤		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
4080	羣	138	138	82FC	7FA3	羣		
4095	舩	7	7	834C	8229	舩		
4107	葱	8	8	8358	8471	葱		
4109	蒭	2	2	835A	84AD	蒭		
4132		221	221	8371			血⿱亼巫	待造字
4134	衞	132	132	8373	885E	衞		
4140	袵	1	1	8379	88B5	袵		
4142	褎	17	17	837B	8943	褎		
4146	覩	10	10	83A1	89A9	覩		
4172	蹙	1	1	83BB	8E77	蹙		Unicode 字型呈現差異
4184	迹	30	30	83C7	8FF9	迹		
4190	郄	4	4	83CD	90C4	郄		
4202	隖	1	1	83D9	9696	隖		
4216	韈	2	2	83E7	97C8	韈		
4237	髡	5	5	83FC	9AE0	髡		
4240	髴	2	2	8440	9AF4	髴		
4242	鬭	30	30	8442	9B2D	鬭		
4266	鼂	16	16	845A	9F02	鼂		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
4293		1	1	8475			○○束⿱火	待造字
4460	勑	1	1	85A1	52D1	勑		
4469	匀	1	1	85AA	5304	匀		
4515	呪	1	1	85D8	546A	呪		
4554	嘧	2	2	8640	55C1	嘧		
4669	轑	7	7	86D5	23358		轑	
4773	姊	1	1	877E	59C9	姊		
4779	鞲	1	1	87A6	97B2	鞲		
4825	鬧	1	1	87D4	28D66		門⿰俞	
4869	孋	4	4	8841	5B4A	孋		
4878	穴	1	1	884A	5B82	穴		
4904	尅	6	6	8864	5C05	尅		
4996	斷	1	1	88E2	65B5	斷		
5065	忼	1	1	8968	5FFC	忼		
5119	慱	2	2	89C0	613D	慱		
5135		1	1	89D0			辶⿰亏	待造字
5169	尼	1	1	89F2	6239	尼		
5182		1	1	8A40			馬⿱衣	待造字

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
5225	卹	6	6	8A6B	460F		血⿰丿卩	
5238		1	1	8A78			⿰形木厂⿱艹⊡	待造字
5244	孌	1	1	8A7E	3738		薛⿱亼女	
5245		1	1	8AA1			掉⿰丿戈	待造字
5289	愨	1	1	8ACD	6164	愨		
5302	昏	8	8	8ADA	662C	昏		
5310	舄	1	1	8AE2	8204	舄		
5317	毟	1	1	8AE9	26248		烝⿱亼缶	
5344	嗀	2	2	8B45	55F8	嗀		
5371	冐	3	3	8B60	808E	冐		
5455	郪	1	1	8BD6	48DB		柰⿰丿卩	
5459	鈌	1	1	8BDA	21647		垂⿰丿夬	
5514	盩	1	1	8C52	25095		執⿱亼皿	
5549	藁	3	3	8C75	85F3	藁		
5564	縻	1	1	8CA6	6ABF	縻		
5579	宓	1	1	8CB5	5D08	宓		
5586	歔	1	1	8CBC	6B57	歔		
5588	惤	1	1	8CBE	6061	惤		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
5651	黿	1	1	8CFD	3E10		難△灬	
5655	𪓐	1	1	8D42	56BB	𪓐		
5697	煑	3	3	8D6C	7151	煑		Unicode 字型呈現差異
5705	廐	3	3	8D74	5ED0	廐		Unicode 字型呈現差異
5706	敍	3	3	8D75	654D	敍		
5717	陾	1	1	8DA2	967F	陾		
5746		2	2	8DBF			每⧋系	待造字
5764	籣	1	1	8DD1	7C5E	籣		
5780	襢	3	3	8DE1	4120		衤⧋亶	
5854	幺	1	1	C6CD	5E7A	幺		
5860	无	2	2	C6D3	65E0		无	

附件二、《前漢紀》「前漢紀.xml」手動取代表

編號	造字	部件	檔案原文摘錄	手動取代結果
2490			封龍__侯	封龍頷侯

附件三、《前漢紀》「前漢紀.xml」「●」字形手動取代表

檔案原文摘錄	手動取代結果
奇●善言	奇⺮△宋善言
嬰沛●騶	嬰沛广△殷騶

遣兵距●關	遣兵距山△堯關
都●災	都广△殷災
蓋六●退飛	蓋六鳥△兒退飛
後宮滿●掌養事數十人	後宮滿广△殷掌養事數十人
乘輿●馬帷帳器物	乘輿广△殷馬帷帳器物
今以虧形在●茸之間	今以虧形在艹△闞茸之間
四面營陣被甲鞮●	四面營陣被甲鞮形牙文革⊡