



《華陽國志校補圖注》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
2009/02/02 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 14 的「胤」字，Unicode 編碼是 38E7，由於 XP 並不支援，仍需使用構字式「彳彡月廾」。

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：



《華陽國志校補圖注》使用舊版造字 115 個，字頻 586 次，這 115 個造字中，101 個可轉成 Windows XP 能顯示的字，字頻 533 次；另外 14 個字必須轉成構字式，字頻 53 次。

轉碼完成製作轉碼分析表，請參考附件一《華陽國志校補圖注》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 1403 的「穴厶兔」字，而原書使用之字為「冤」，查詢教育部異體字字典確認「穴厶兔」並非「冤」之異體字，因此歸為錯字，以程式取代為正確字形「冤」。
 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. 本無該字，手動刪除：檔案中使用的舊版造字為空白字形，校對原書時也無該字，手動刪除。例：造字編號 86 的舊版造字為空白字形，原書也無該字，手動刪除。手動修改部份請參考附件二《華陽國志校補圖注》手動取代表。

附件一、《華陽國志校補圖注》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
14	胤	3	FA4D	38E7		 彳 幺 月 乚 	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
86		1	FAB7				本無該字，手動刪除。
821	脊	1	8E63	7718	脊		
842	蚪	11	8E78	3AB7		蚪	
893	峩	2	8ECD	5CE9	峩		
982	胷	1	8F67	80F7	胷		
1024	虵	5	8FB3	8675	虵		
1039	稟	1	8FC2	69C0	稟		
1067	碁	1	8FDE	68CA	碁		
1073		1	8FE4			算会么	篡，異體字問題。
1077	稟	1	8FE8	7A3E	稟		
1079	總	20	8FEA	7DEB	總		
1111	瓚	7	904B	749D	瓚		
1192	盥	1	90BE	76E2	盥		
1403		2	91F4			穴会免	冤，錯字，以程式全部取代。
1513	糧	1	92C5	7CAE	糧		糧，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
1530	紉	1	92D6	7CFA	紉		
1565	繇	52	92F9	7DDC	繇		
1614	聳	1	936B	8053	聳		
1642	脉	2	93A9	8109	脉		
2179	胷	1	96EB	26643		胷	
2254	喰	1	9777	20DE3		口 𠂇 虫	
2386	瑯	1	985E	7458	瑯		
2830	眈	3	9B43	4022		目 𠂇 矢	
2878	叟	9	9B73	5081	叟		
2947	聰	4	9BDA	8066	聰		
3780	儁	11	814B	5101	儁		
3789	冲	3	8154	51B2	冲		沖，異體字問題。
3791	况	8	8156	51B5	况		況，異體字問題。
3796	决	1	815B	51B3	决		決，異體字問題。
3804	効	10	8163	52B9	効		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3805	勅	5	8164	52C5	勅		
3814	却	6	816D	5374	却		
3830	咤	1	817D	54A4	咤		
3840	坂	9	81A9	5742	坂		
3846	壻	3	81AF	58FB	壻		
3850	姪	3	81B3	59D9	姪		
3858	寃	4	81BB	5BC3	寃		寃，異體字問題。
3864	峯	1	81C1	5CEF	峯		
3869	帀	1	81C6	5E00	帀		
3876	廻	1	81CD	5EFB	廻		
3884	忽	2	81D5	6031	忽		
3891	慙	6	81DC	6159	慙		
3924	朞	6	81FD	671E	朞		
3930	挀	3	8244	6901	挀		
3961	灾	5	8263	707E	灾		
3974	牀	2	8270	7240	牀		
3976	犁	2	8272	7282	犁		
3978	猪	7	8274	732A	猪		Unicode 字型

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							呈現差異。
3980	獾	3	8276	7328	獾		
3984	瑠	1	827A	7460	瑠		
3992	嘗	3	82A4	751E	嘗		
3996	疎	5	82A8	758E	疎		
4015	眦	1	82BB	7726	眦		
4041	豎	1	82D5	7AEA	豎		
4080	羣	43	82FC	7FA3	羣		
4081	翱	1	82FD	7FFA	翱		
4089	脇	2	8346	8107	脇		
4091	臯	2	8348	81EF	臯		
4095	舩	5	834C	8229	舩		
4104	莅	6	8355	8385	莅		
4107	葱	1	8358	8471	葱		
4109	葛	1	835A	84AD	葛		
4132		108	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	32	8373	885E	衛		衛，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4138	裘	1	8377	88A0	裘		
4139	袴	1	8378	88B4	袴		
4146	覩	4	83A1	89A9	覩		睹，異體字問題。
4169	趨	1	83B8	8DA6	趨		
4173	駝	6	83BC	8EAD	駝		
4174	駱	5	83BD	8EB6	駱		
4184	迹	2	83C7	8FF9	迹		
4185	邊	3	83C8	9037	邊		
4190	邳	7	83CD	90C4	邳		
4198	鑛	5	83D5	945B	鑛		
4202	隄	1	83D9	9696	隄		
4240	鬚	1	8440	9AF4	鬚		
4242	鬪	2	8442	9B2D	鬪		
4258	麇	1	8452	9E90	麇		
4262	麪	1	8456	9EAA	麪		麪，異體字問題。
4267	鼈	1	845B	9F08	鼈		
4275	义	2	8463	4E49	义		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4309	刼	5	84A7	523C	刼		
4414	幾	1	8551	27C19		豈厶幾	
4455	劒	2	857A			僉厶刃	劒，異體字問題。
4460	勅	12	85A1	52D1	勅		
4639	駟	1	86B7	48D6		馬厶邑	
4655	垠	1	86C7	57BB	垠		
4723	壇	1	874C	58C3	壇		
4768	妬	1	8779	59AC	妬		妒，異體字問題。
4827	𦉳	2	87D6	3DC0		〇火厶 𦉳	
4857	嶸	3	87F4	21E2B		山厶并	
4862	柳	2	87F9	678A	柳		
4904	尅	5	8864	5C05	尅		
4916	殺	1	8870	27623		宀厶殳	
4996	斲	2	88E2	65B5	斲		
5012	媿	2	88F2	36D0		女厶更	
5034	弃	1	8949	5F03	弃		棄，異體字問題

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
5116	侷	3	89BD	4F37	侷		
5133	憇	1	89CE	6199	憇		
5192	隙	2	8A4A	969F	隙		
5220	𠃉	1	8A66	6BCC	𠃉		
5268	遲	1	8AB8	28488		辵 彳 尸	
5300	鬪	1	8AD8	9B2A	鬪		鬪，異體字問題。
5450	滄	1	8BD1	6E4C	滄		
5463	𡗗	1	8BDE	65F9	𡗗		
5466	倏	1	8BE1	5010	倏		
5578	弛	2	8CB4	38AE		弓 彡 屯	
5588	恡	1	8CBE	6061	恡		
5655	蹶	4	8D42	56BB	蹶		
5678	𡗗	1	8D59	757A	𡗗		
5697	𡗗	3	8D6C	7151	𡗗		
5706	敘	5	8D75	654D	敘		敘，異體字問題。
5709	蟀	2	8D78	872F	蟀		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5723	样	21	8DA8	4367		牛 𠂔 羊	

二、手動取代表說明：

《華陽國志校補圖注》檔案中需手動修改的情況分為以下幾種：

- (一) 《華陽國志校補圖注》檔案中之舊版造字，因字形部份空白，原書也無此字形，手動刪除，並製作手動取代表，請參考附件二。欄位說明如下：

1. 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
2. 造字：舊版造字字形。
3. 檔案原文摘錄：摘錄檔案《華陽國志校補圖注》中的文句段落，其中紅字底線的部分，為需要手動取代的原文。
4. 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。
5. 備註：記錄修改相同辭句的次數或其他事項。

附件二、《華陽國志校補圖注》手動取代表

編號	造字	檔案原文摘錄	手動取代結果	備註
86		六十三人士，十一人女 <u> </u> 。	六十三人士，十一人女·	

- (二) 《華陽國志校補圖注》檔案中出現「●」字形，共計 39 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，請參考附件三。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《華陽國志校補圖注》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。

附件三、《華陽國志校補圖注》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 5	給客橙、●	蔘	
p. 14	今所謂●頭虎子者	弜	
p. 30	謝、然、●、楊	忛 厶蓋	
p. 36	有●頭白虎復	弜	
p. 79	清水出●	魚 厶奧	
p. 163	●\Qa 諷、尹睦	役	
p. 209	照●則使人	面	
p. 240	●腦三蚪	蟠	
p. 259	●\Qa 山為田	畚	
p. 261	郡治·有●津	石 厶兆	
p. 285	●溪，置嵩唐	汜 厶耆	
p. 346	【諛】\qa●\Qa	設	
p. 429	籌斥●元豐	𦰩 厶下	
p. 429	廖本作●	𦰩 厶下	
p. 445	蟀●羌邳逢	虫 厶同	
p. 537	李●，字孟元	纘 厶几	
p. 537	●固爭之	纘 厶几	

頁碼	檔案原文摘錄	原書字形	備註
p. 537	●對以“詔書	纘𠄎几	
p. 537	●掾	纘𠄎几	
p. 537	仲【●】\qa 昱\Qa	日𠄎工	待造字 有5處
p. 592	進楊●●	𠄎禾白ナッ𠄎 𠄎禾白ナッ𠄎	
p. 592	閨門雍●	𠄎禾白ナッ𠄎	
p. 592	●封未定	𠄎フ白丨𠄎	
p. 593	【●】(玦)	王𠄎央	待造字 有5處
p. 593	【●】(蜀)穀二石	𠄎𠄎𠄎𠄎𠄎	待造字
p. 616	姜●康、稚	來𠄎女	
p. 664	【●】\qa 琰\Qa	𠄎王火人𠄎	待造字
p. 669	司空辟士李●	纘𠄎几	
p. 669	節士，仲【●】	日𠄎工	待造字
p. 699	程貞【●】	王𠄎央	待造字

(三)《華陽國志校補圖注》檔案中，對於舊版造字以外發現的錯字，以手動修改，並製作錯字修改表，請參考附件四。欄位說明如下：

1. 檔案原文摘錄：摘錄檔案《華陽國志校補圖注》中的文句段落，其中紅字底線的部分，為需要手動取代的原文。
2. 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。

3. 備註：記錄修改相同詞句的次數或其他事項。

附件四、《華陽國志校補圖注》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
尅捷籌 斥 ●元豐	尅捷籌●元豐	