

《鹽鐵論》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
2010/06/04 丁玫伶

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 2160 的「磳」字，Unicode 編碼是 40D8，由於 XP 並不支援，仍需使用構字式「石𠩺堅」。

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：

《鹽鐵論》使用舊版造字 69 個，字頻 152 次，這 69 個造字中，58 個可轉成 Windows XP 能顯示的字，字頻 139 次；另外 11 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《鹽鐵論》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 2449 的「雒」字，而原書使用之字為「頡」，查詢教育部異體字字典確認「頡」並非「雒」之異體字，因此歸為錯字，以程式取代為正確字形「頡」。
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 4091 的「臯」字，Unicode 字型呈現為「臯」，實際上仍為同一字。

附件一、《鹽鐵論》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
827	鮫	1	8E69	9B8C	鮫		
845	鼯	2	8E7B	9F26	鼯		鼠  召，異體字問題。
881	颺	1	8EC1	98C7	颺		
899	駝	2	8ED3	99DE	駝		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1024	虵	2	8FB3	8675	虵		
1044	儻	2	8FC7	511B	儻		
1226	罰	4	90E0	7F78	罰		
1689	禱	1	93D8	464F		礻 ㄖ 專	
1981	讐	1	95C2	8B90	讐		
2160	磴	1	96D8	40D8		石 ㄖ 堅	
2449	雒	1	98BF			名 ㄖ 隹	額，錯字，以程式全部取代。
2524	饒	1	994B	994D	饒		
2659	瓜	1	99F4	201D7		亻 ㄖ 瓜	
3794	湊	1	8159	51D1	湊		
3798	𠂇	1	815D	51F7	𠂇		
3804	効	3	8163	52B9	効		
3840	坂	3	81A9	5742	坂		
3868	卮	1	81C5	5DF5	卮		
3883	徧	2	81D4	5FA7	徧		
3891	慙	1	81DC	6159	慙		
3894	筋	2	81DF	4225		竹 ㄖ 勛	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3930	惇	2	8244	6901	惇		
3937	檝	2	824B	6A9D	檝		
3974	牀	3	8270	7240	牀		
3984	瑠	1	827A	7460	瑠		
4010	臯	2	82B6	7690	臯		臯，異體字問題。
4014	蓋	2	82BA	8462	蓋		
4035	穽	3	82CF	7A7D	穽		
4077	罇	1	82F9	7F47	罇		
4085	耻	7	8342	803B	耻		
4091	臯	4	8348	81EF	臯		Unicode 字型呈現差異。
4095	舩	1	834C	8229	舩		
4134	衛	10	8373	885E	衛		衛，異體字問題。
4139	袴	2	8378	88B4	袴		
4140	袵	1	8379	88B5	袵		
4146	覩	25	83A1	89A9	覩		
4150	訕	2	83A5	8A29	訕		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4172	蹙	1	83BB	8E77	蹙		
4184	迹	3	83C7	8FF9	迹		
4222	頰	1	83ED	9847	頰		
4226	凜	2	83F1	98E1	凜		
4234	羸	2	83F9	9A58	羸		
4242	鬪	3	8442	9B2D	鬪		
4265	鼃	3	8459	9F03	鼃		
4267	鼈	2	845B	9F08	鼈		
4458	勑	2	857D	52CC	勑		
4558	廐	1	8644	53A9	廐		廐，異體字問題。
4561		1	8647			廐 ^廐 廐 廐比廐	廐，錯字，以程式全部取代。
4727	壘	1	8750	58C4	壘		
4768	妬	5	8779	59AC	妬		
4825	闢	1	87D4	28D66		門△俞	
4887	醞	1	8853	4908		酉△益	
4910		1	886A			氵△畷	淫，錯字，以

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							程式全部取代。
4981	鳧	1	88D3	9CEC	鳧		鳧，異體字問題。
4996	斲	1	88E2	65B5	斲		
5045	飪	1	8954	9901	飪		
5087	躡	1	897E	2810C		足 厶 習	
5265	蓄	1	8AB5	26E1C		艹 厶 艸 厶 回	
5308		2	8AE0			棹 厶 戈	檝，錯字，以程式全部取代。
5434	湟	1	8BC1	6E7C	湟		
5466	倏	1	8BE1	5010	倏		
5497	松	1	8C41	8873	松		
5588	恠	1	8CBE	6061	恠		
5722	朞	2	8DA7	671E	朞		
5723	牂	1	8DA8	4367		牛 厶 羊	牂，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5724	稽	2	8DA9	7A38	稽		
5746	繇	1	8DBF			每厶系	
5747	瘠	1	8DC0	3FCA		疒厶替	
5764	籟	2	8DD1	7C5E	籟		

二、手動取代表說明：

《鹽鐵論》檔案中出現「●」字形，共計17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《鹽鐵論》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、《鹽鐵論》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 190	數徙則●	歹厶委	
p. 351	鮮羔●，幾胎	羊厶兆	
p. 351	豐●耳菜	艹厶粟	
p. 351	堊●壁飾	巾厶慶	
p. 352	單●木具	復	

頁碼	檔案原文摘錄	原書字形	備註
p. 352	燕●代黃	鼠厶各	
p. 352	富者●耳	革厶真	
p. 352	耳銀鑷●	革厶巛	
p. 352	華●明鮮	革厶茸	
p. 354	中者●皮	犛厶莫	待造字
p. 354	狗●馬腹	月厶習	
p. 473	糠●之苦	米厶后	
p. 473	老母之●頤	顛	
p. 548	犀●兕甲	革厶由	
p. 548	以道德為●	革厶由	
p. 592	●裂可以補	巾厶祭	
p. 629	切而不●	火厶索	