

《淮南子》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
2010/05/14 丁玟伶

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 1197 的「礪」字，Unicode 編碼是 40FA，由於 XP 並不支援，仍需使用構字式「石𠩺靡」。

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：

《淮南子》使用舊版造字 142 個，字頻 466 次，這 142 個造字中，121 個可轉成 Windows XP 能顯示的字，字頻 417 次；另外 21 個字必須轉成構字式，字頻 49 次。

轉碼完成製作轉碼分析表，請參考附件一《淮南子》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. 部份使用不同字形，手動修改：檔案中為同一舊版造字，校對時卻發現原書使用兩個以上不同的字形，可能部份舊版造字為錯字或原書使用兩種異體字形，導致同一舊版造字轉碼時對應到兩個以上的字形；將一部份手動修改，剩餘的頻次則用程式取代成對應的 Unicode 或構字式。例：造字編號 4878 的「宄」字頻次為 2，依書將 1 字手動修改為「宄」，剩餘 1 字轉為「宄」。手動修改部份請參考附件二《淮南子》手動取代表。
 3. 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 936 的「義𪛗鳥」字，而原書使用之字為「壽𪛗鳥」，查詢教育部異體字字典確認「壽𪛗鳥」並非「義𪛗鳥」之異體字，因此歸為錯字，以程式取代為正確字形「壽𪛗鳥」。
 4. 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 3935 的「槩」字，Unicode 字型呈現為「槩」，實際上仍為同一字。

附件一、《淮南子》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
----	----	----	------	---------	-------	-----	----

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
814	墓	2	8E5C	87C7	墓		
841	礮	1	8E77	7932	礮		
852	貉	1	8EA4	72E2	貉		
869	糜	1	8EB5	6A5C	糜		
895	烽	1	8ECF	71A2	烽		
896	櫛	1	8ED0	6AE9	櫛		
936	鷓	2	8EF8	4D0A		義𪗇鳥	壽𪗇鳥，錯字，以程式全部取代。
957	遨	2	8F4E	906C	遨		
982	胷	11	8F67	80F7	胷		
1067	碁	4	8FDE	68CA	碁		
1077	稟	1	8FE8	7A3E	稟		
1081	鶴	1	8FEC	9E16	鶴		
1088	暘	1	8FF3	757C	暘		
1091	漣	3	8FF6	58CD	漣		
1133	瓌	2	9061	74CC	瓌		
1197	礪	2	90C3	40FA		石𪗇靡	
1205	眎	1	90CB	770E	眎		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1239	豎	2	90ED			豎 𠄎豆	豎，異體字問題。
1565	絲	3	92F9	7DDC	絲		
1642	脉	3	93A9	8109	脉		
1843	擘	1	94D5	22BF2		擘 𠄎手	
1871	蟲	4	94F1	87C1	蟲		
1912	襍	1	955B	894D	襍		
1937	詢	1	9574	8A3D	詢		
1981	讐	4	95C2	8B90	讐		
2050	蹠	13	9648	8E4F	蹠		
2077	鞞	1	9663	8F4A	鞞		
2162	醜	1	96DA	918E	醜		
2163	桴	1	96DB	7CB0	桴		
2195	攢	1	96FB	3A4C		攢 𠄎費	
2416	筮	1	987C	25AF1		筮 𠄎屯	
2445	鞞	1	98BB	2930A		鞞 𠄎革	鞞，錯字，以程式全部取代。
2572	銜	1	997B	8858	銜		銜，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
2610	鱣	5	99C3	9C53	鱣		
2638	鷓	2	99DF	9D55	鷓		
2783	蚤	2	9AD3	8689	蚤		
2817	窈	13	9AF5	3AC4		窈	
2871	穎	1	9B6C	8D11	穎		
3016	闕	1	9C60	28D91		門△臭	闕，異體字問題。
3017	圃	1	9C61	4A49		面△甫	
3770	个	7	8141	4E2A	个		
3780	儻	1	814B	5101	儻		
3791	況	1	8156	51B5	況		況，異體字問題。
3796	決	1	815B	51B3	決		決，異體字問題。
3804	効	4	8163	52B9	効		
3814	却	3	816D	5374	却		
3819	廡	1	8172	53AE	廡		廡，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3820	厦	1	8173	53A6	厦		
3834	嗶	2	81A3	5637	嗶		嗶，異體字問題。
3868	卮	4	81C5	5DF5	卮		
3869	币	2	81C6	5E00	币		
3880	護	1	81D1	5F60	護		
3883	徧	8	81D4	5FA7	徧		
3891	慙	5	81DC	6159	慙		
3894	筋	3	81DF	4225		竹 𠂇 勛	
3915	斃	1	8459	6585	斃		
3924	菴	2	81FD	671E	菴		
3930	憵	1	8244	6901	憵		
3935	槩	6	8249	3BA3	槩		Unicode 字型呈現差異。
3937	檝	3	824B	6A9D	檝		
3955	澗	1	825D	3D4E		氵 𠂇 間	澗，異體字問題。
3966	烟	1	8268	70DF	烟		
3971	熇	1	826D	71BA	熇		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3974	牀	4	8270	7240	牀		
3976	犁	1	8272	7282	犁		
3980	獾	8	8276	7328	獾		
3996	疎	2	82A8	758E	疎		
4010	皐	5	82B6	7690	皐		皐，異體字問題。
4011	皞	2	82B7	76A1	皞		皞，異體字問題。
4015	眦	2	82BB	7726	眦		
4017	螻	1	82BD	880F	螻		
4034	穉	1	82CE	7A49	穉		
4035	穽	1	82CF	7A7D	穽		
4038	竈	8	82D2	7AC8	竈		
4040	竝	23	82D4	7ADD	竝		
4045	箒	1	82D9	7B92	箒		
4051	籊	1	82DF	7C35	籊		
4066	綺	1	82EE	7D5D	綺		
4068	綫	1	82F0	7DAB	綫		
4085	耻	6	8342	803B	耻		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4089	脇	3	8346	8107	脇		
4091	臯	3	8348	81EF	臯		Unicode 字型呈現差異。
4100	苳	2	8351	82FD	苳		
4115	藁	1	8360	85C1	藁		
4124	蝨	4	8369	8771	蝨		
4129	虬	2	836E	866C	虬		
4134	衛	25	8373	885E	衛		衛，異體字問題。
4142	裛	1	837B	8943	裛		
4145	羈	1	837E	898A	羈		Unicode 字型呈現差異。
4146	覩	6	83A1	89A9	覩		
4155	訶	9	83AA	8B0C	訶		
4160	適	1	83AF	8B81	適		
4165	獾	1	83B4	8C9B	獾		
4172	屨	8	83BB	8E77	屨		部份使用不同字形「屨」和「楚」，手動修

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							改。
4184	迹	35	83C7	8FF9	迹		
4193	鈎	1	83D0	920E	鈎		
4222	頰	1	83ED	9847	頰		
4225	顛	1	83F0	9866	顛		
4230	聊	1	83F5	99E0	聊		
4242	鬪	29	8442	9B2D	鬪		
4252	鰓	2	844C			魚厶眾	鰓，異體字問題。
4267	鼈	12	845B	9F08	鼈		
4336	𧢲	1	84C2	278BB		角厶牛	
4342	倮	5	84C8	502E	倮		
4437	廩	1	8568	5EEA	廩		
4454	𠂔	2	8579	2595B		穴厶友	
4455	劒	2	857A			僉厶刀	
4502	俎	1	85CB	723C	俎		
4533	唵	1	85EA	552B	唵		
4554	唬	2	8640	55C1	唬		
4707	倭	1	86FB	3944		亻厶夊	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4735	壠	2	8758	58E0	壠		
4744	絀	2	8761	25FC9		糸 厶 末	
4749	鶯	1	8766	2A1AA		殼 厶 鳥	
4768	妬	1	8779	59AC	妬		
4794	藁	1	87B5	8448	藁		
4843	昏	1	87E6	4055		目 厶 昏	
4878	宄	2	884A	5B82	宄		部份使用不同字形「宄」，手動修改。 Unicode 字型呈現差異。
4904	尅	3	8864	5C05	尅		
4996	斲	16	88E2	65B5	斲		
5019	畧	1	88F9			畧	
5034	弃	2	8949	5F03	弃		棄，異體字問題。
5043	曠	1	8952	5F4D	曠		
5074	尫	1	8971	5C2B	尫		
5108	惛	4	89B5	60FD	惛		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5147	恠	1	89DC	6060	恠		
5302	昏	4	8ADA	662C	昏		
5335	𧈧	5	8AFB	459F		亡厶虫	
5396	鎔	1	8B79	938B	鎔		
5438	𦉑	1	8BC5	681E	𦉑		
5439	栢	2	8BC6	6822	栢		
5461	𡗗	2	8BDC	6B6B	𡗗		
5464	煥	4	8BDF	7157	煥		
5473	攷	1	8BE8	22BF1		扌厶寔	
5578	弛	2	8CB4	38AE		弓厶龟	
5582	鞣	1	8CB8	97B5	鞣		
5583	殮	1	8CB9	98F1	殮		
5651	難	1	8CFD	3E10		難厶灬	
5705	廐	1	8D74	5ED0	廐		Unicode 字型 呈現差異。
5722	朞	1	8DA7	671E	朞		
5724	稽	1	8DA9	7A38	稽		
5783	教	5	8DE4	3A8D		孛厶攴	

二、手動取代表說明：

《淮南子》檔案中需手動修改的情況分為以下幾種：

(一)《淮南子》檔案中之舊版造字，因字形部份不同而無法全部以程式取代，皆手動修改成Windows XP能顯示的字或構字式，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. 取代：實際使用字形與舊版造字不同時，於此欄列出。
4. 檔案原文摘錄：摘錄檔案《淮南子》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。
5. 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。
6. 備註：記錄修改相同辭句的次數或其他事項。

附件二、《淮南子》手動取代表

編號	造字	取代	檔案原文摘錄	手動取代結果	備註
4172	𧇗	𧇗	則後者 <u>𧇗</u> 之	則後者 <u>展△足</u> 之	
		𧇗	舉足 <u>𧇗</u> 其體	舉足 <u>𧇗</u> 其體	有 2 處
4878	宀	宀	纏錦經 <u>𧇗</u>	纏錦經 <u>宀</u>	

(二)《淮南子》檔案中出現「☆」字形，共計 125 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《淮南子》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、《淮南子》「☆」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 6	☆遠彌高以	歷	
p. 20	☆渡汶而死	𧇗	
p. 20	偶☆智故	目𠂔差	
p. 32	堅強而不☆	鞮	
p. 40	☆然能聽	𦉳𠂔目	
p. 52	鈞而不☆	石𠂔為	
p. 56	☆，皆本於	𠂔木人羊𠂔	
p. 57	無☆而復反	𠂔來女土𠂔	
p. 59	☆，雜之以青黃	屈𠂔𠂔	
p. 65	昧昧☆☆	𦉳𦉳	
p. 65	重九☆	𦉳	
p. 67	握德☆性	𠂔扌寒巳𠂔	
p. 73	疾☆、辟寒暑	羊𠂔虫	
p. 73	蚤蟲☆膚	口𠂔替	
p. 74	則必無餘☆	𠂔木人羊𠂔	
p. 78	☆罟張而在下	𠂔𠂔亡	
p. 142	多力而☆	𠂔𠂔犬𠂔犬𠂔 犬𠂔	待造字
p. 145	大口決☆	眦	

頁碼	檔案原文摘錄	原書字形	備註
p. 155	玄☆	龜	有 2 處
p. 156	凡根☆草者	芟	
p. 156	不根☆者	芟	
p. 165	置罌羅☆	罌 罌	
p. 166	乃合☆牛	牛 𠂔 累	
p. 171	腐草化為☆	蚘	
p. 184	介蟲為☆	介 𠂔 芟	
p. 201	而甘☆	卧	
p. 207	☆倨倨	卧	
p. 210	社槁而☆裂	土 𠂔 零	
p. 215	不☆喋苛事	口 𠂔 嬰	
p. 228	猶☆醜也	其 𠂔 頁	
p. 233	☆下迫頤	月 𠂔 曷	
p. . 238	與守其☆	篇	
p. 244	山石，☆金玉	金 𠂔 挈	
p. 247	剝☆削鋸	屈 𠂔 𠂔	
p. 253	餘糧於☆首	晦	
p. 260	剛而不☆	鞮	
p. 260	則☆隘而不容	陘	
p. 273	孫叔敖恬☆	卧	
p. 295	夫寸生於☆	禾 𠂔 粟	

頁碼	檔案原文摘錄	原書字形	備註
p. 295	☆生於日	禾厶粟	
p. 300	馬☆于衡下	夕厶人	改為死
p. 318	其寢無☆	癡	
p. 330	理詘☆倨尙	亻厶曳	
p. 331	宿諸☆首	晦	
p. 339	大絃☆	絃	
p. 341	天地能☆之	函	
p. 345	水蠶為☆☆	形虫每心口蕊	
p. 346	丘生乎象☆	木厶著	待造字
p. 349	角☆不厭薄	角厶喬	
p. 359	壤土草☆而已	形艹殺 卩口	
p. 373	☆胡飲水	鳥厶弟	
p. 407	岡☆之野	宀厶良	
p. 421	不勝暑熱☆	民厶虫	
p. 425	知☆護之	桀	
p. 446	☆寸而伸尺	誦	
p. 458	枕戶隣而☆	卧	有 2 處
p. 462	☆足眾而走	蚘	
p. 482	匹夫百☆一守	晦	
p. 501	不見朕☆	畎厶土	
p. 503	若☆之足	蚘	有 2 處

頁碼	檔案原文摘錄	原書字形	備註
p. 511	☆☆如夏	日 𠄎 廣 日 𠄎 廣	
p. 516	☆巨旗	扌 𠄎 擧	
p. 529	☆以為鼓	鞞	
p. 543	犁牛，既☆	耨	
p. 543	以☆，決鼻	牛 𠄎 脩	
p. 546	稅其☆	車 𠄎 義	
p. 546	在☆茵之上	衤 𠄎 丹	待造字
p. 547	☆伯玉以德	遽	
p. 548	雞頭已☆	瘻	
p. 550	足☆地而為	展 𠄎 足	
p. 553	☆靡勿釋	幾 𠄎 𠄎	
p. 555	足以☆者	展 𠄎 足	
p. 555	不☆而後行	展 𠄎 足	
p. 555	以手☆	扌 𠄎 市	
p. 557	而☆皆如一	扌 𠄎 弃	
p. 557	日出☆谷	暘	
p. 558	羊去蚤☆	𠄎 𠄎 虫	
p. 559	則☆履之	展 𠄎 足	
p. 562	射者遺其☆	執	
p. 564	☆苗類絮	蒿	

頁碼	檔案原文摘錄	原書字形	備註
p. 570	蟣☆相弔	孔𠄎虫	
p. 571	☆者扣舟	𠄎𠄎瓜	
p. 572	頭☆與空木	孔𠄎虫	
p. 576	若☆薄冰蛟	展𠄎足	
p. 576	☆面𠄎甫在頰	面𠄎馱	
p. 576	始於耒☆	耜	
p. 577	子子為☆	民𠄎虫	
p. 578	☆燭掬	膺	
p. 584	賊心☆	亡𠄎虫	
p. 586	前而不☆	執𠄎車	
p. 586	誹☆己者	謗	
p. 591	☆虐萬民	暴	
p. 602	☆不能蔽也	暴	
p. 608	人之思☆	歿	
p. 608	征☆亂	暴	
p. 615	一指☆之	𠄎𠄎𠄎𠄎伐	
p. 620	☆亂之所致	暴	
p. 633	孔子無黔☆	突	
p. 639	沉☆耽荒	酉𠄎面	
p. 639	☆☆哆嗚	嗚月𠄎癸	
p. 642	欲☆水也	展𠄎足	

頁碼	檔案原文摘錄	原書字形	備註
p. 643	摩其鋒☆	𠄎𠄎 𠄎	
p. 647	故弓待☆	木𠄎 敬	
p. 652	☆，夫婦男女	𠄎 月	
p. 659	☆面𠄎甫搖	面𠄎 獸	
p. 659	便☆擬神	媚	
p. 667	蟻☆	蟲	有 2 處
p. 669	讓長，而☆白	辨	
p. 676	☆虛輕舉	足𠄎 奈	
p. 678	有☆，秤薪	𠄎 是	
p. 680	貴其☆息脉血	摩	
p. 698	百☆，樹一	隙	
p. 698	昌羊去蚤☆	蟲	
p. 700	靡☆大宗	散	
p. 704	齟之☆也	邈	
p. 707	而☆逐萬物	形 扌 鹿 君 口	
p. 708	卑弱制強☆	暴	

(三)《淮南子》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《淮南子》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、《淮南子》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
被之而不 <u>裛</u>	被之而不 <u>褻</u>	
出，☆ <u>醜</u> 搖	出，☆ <u>面</u> <u>▲</u> <u>甫</u> 搖	