

《呂氏春秋》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
2008/4/24 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 1843 的「擊」字，Unicode 編碼是 22BF2，由於 XP 並不支援，仍需使用構字式「𠄎𠄎手」。

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：

《呂氏春秋》使用舊版造字 76 個，字頻 313 次，這 76 個造字中，67 個可轉成 Windows XP 能顯示的字，字頻 304 次；另外 9 個字必須轉成構字式，字頻 9 次。

轉碼完成製作轉碼分析表，請參考附件一《呂氏春秋》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 2559 沒有對應之字，而原書使用之字為「弛」，所以歸為錯字，以程式取代為正確字形「弛」。
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 5381 的「脛」字，Unicode 字型呈現為「脛」，實際上仍為同一字。

附件一、《呂氏春秋》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
827	鮫	3	8E69	9B8C	鮫		
957	遨	3	8F4E	906C	遨		
972	犇	1	8F5D	7287	犇		
982	胃	4	8F67	80F7	胃		胸，異體字問題。
1024	虵	1	8FB3	8675	虵		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1053	傲	1	8FD0	6160	傲		
1062		4	8FD9			馬厶亞	駢，異體字問題。
1099	睺	5	8FFE	66A4	睺		
1133	瓌	1	9061	74CC	瓌		
1243	桐	4	90F1	79F1	桐		
1843	擘	1	94D5	22BF2		収𠂇手	
1857	蚕	1	94E3	86A0	蚕		
1885	蠶	2	9540	882D	蠶		
1937	詢	2	9574	8A3D	詢		
2053	躡	1	964B	8EA7	躡		
2146	蝸	1	96CA	8739	蝸		
2162	馘	1	96DA	918E	馘		
2559		1	996E				弛，錯字，以程式全部取代。
2610	鯁	2	99C3	9C53	鯁		
2612	暗	1	99C5	3B10		日厶晉	
3770	个	8	8141	4E2A	个		
3780	雋	2	814B	5101	雋		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3801	剗	3	8160	5257	剗		
3804	効	2	8163	52B9	効		
3814	却	10	816D	5374	却		
3835	噍	1	81A4	564D	噍		
3852	嫵	1	81B5	5A63	嫵		
3855	擘	6	81B8	5B7C	擘		
3873	廢	1	81CA			广△段	廢，異體字問題。
3883	徧	3	81D4	5FA7	徧		
3891	慙	8	81DC	6159	慙		
3924	朞	4	81FD	671E	朞		期，異體字問題。
3974	牀	2	8270	7240	牀		
3980	獯	2	8276	7328	獯		
3995	晦	2	82A7	756E	晦		
4038	竈	9	82D2	7AC8	竈		
4040	竝	10	82D4	7ADD	竝		
4055	粃	7	82E3	7C83	粃		
4072	縲	1	82F4	7DE5	縲		
4104	莅	1	8355	8385	莅		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4115	藁	2	8360	85C1	藁		
4134	衛	76	8373	885E	衛		衛，異體字問題。
4139	袴	2	8378	88B4	袴		
4142	裛	4	837B	8943	裛		裛，異體字問題。
4146	覩	3	83A1	89A9	覩		睹，異體字問題。
4164	凜	1	83B3	23F6C		冫 凵 寒	
4165	獾	2	83B4	8C9B	獾		
4172	麘	6	83BB	8E77	麘		
4180	臯	2	83C3	8FA0	臯		
4184	迹	4	83C7	8FF9	迹		
4190	邳	13	83CD	90C4	邳		
4217	鞞	2	83E8	97E0	鞞		
4242	鬪	31	8442	9B2D	鬪		
4250	鱒	1	844A	9C0C	鱒		
4267	鼈	6	845B	9F08	鼈		
4342	倮	1	84C8	502E	倮		
4533	唸	1	85EA	552B	唸		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4629	埤	2	86AD	577F	埤		
4768	妬	2	8779	59AC	妬		妒，異體字問題。
4822	尫	2	87D1	5C29	尫		
5021	簾	1	88FB	25D27		竹 𠂇 𠂇 𠂇	
5057	搯	1	8960	39FB		扌 𠂇 豕	
5091	鄒	1	89A4	92E3	鄒		
5093	鴿	1	89A6	9D2A	鴿		鴿，異體字問題。
5108	惛	1	89B5	60FD	惛		
5302	昏	2	8ADA	662C	昏		
5306		3	8ADE			竹 𠂇 間	簡，異體字問題。
5364	𦓐	1	8B59	239D1		歹 𠂇 凶	
5371	𦓐	3	8B60	808E	𦓐		
5381	脛	1	8B6A	8103	脛		Unicode 字型呈現差異。
5392	息	2	8B75	60A4	息		
5464	煥	4	8BDF	7157	煥		
5514	𦓐	1	8C52	25095		執 𠂇 皿	
5730	隄	3	8DAF	9638	隄		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5754	躑	1	8DC7	2818C		足厶煩	
5762	截	1	8DCF	229F5		截	

二、手動取代表說明：

《呂氏春秋》檔案中需手動修改的情況分為以下幾種：

(一)《呂氏春秋》檔案中出現「☆」字形，共計 46 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《呂氏春秋》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、《呂氏春秋》「☆」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 55	墨者有鉅子腹☆	黃厶享	
p. 56	腹☆對曰	黃厶享	
p. 56	腹☆不可不行	黃厶享	
p. 96	禽滑☆學於墨子	毆厶康	
p. 96	學於禽滑☆	毆厶康	
p. 136	處目則為☆為盲	目厶蔑	
p. 137	重水所多☆	允厶重	改為樞
p. 285	乃以麋☆置缶	革厶各	
p. 311	腐草化為☆	蚘	

頁碼	檔案原文摘錄	原書字形	備註
p. 359	長短頡☆	吾 厶 午	
p. 402	壯佼老幼胎☆之死者	月 厶 賣	
p. 479	☆其手	磨 厶 卩	
p. 490	過☆涿視	册 厶 頁	
p. 507	刃若新☆	磨 厶 卩	
p. 536	釜☆者而逐之	鬲 厶 瓦	
p. 623	結☆	冈 厶 不	改為罍
p. 623	☆	冈 厶 亡	改為𠂔 厶 亡
p. 628	鄭人之下☆也	革 厶 處	
p. 678	水雲角☆	角 厶 妻 厶 才	
p. 741	☆。肉之美者	月 厶 侯	
p. 741	洞庭之☆	魚 厶 專	
p. 741	陽☆之薑	櫟	
p. 968	無由接而言見，☆	詭	
P. 1065	南撫多☆	嬰 厶 頁	
p. 1156	有執蹠☆而上	疒 厶 裕	
p. 1167	口☆不言	口 厶 昏	
p. 1211	魏令孟印割絳、☆	彤 穴 十 分 口	
p. 1211	以絳、☆、安邑	彤 穴 十 分 口	
p. 1235	丹績之☆	禘 厶 旬	
p. 1293	☆為煩辱	木 厶 異	
p. 1414	許鄙相☆	月 厶 尻	
p. 1414	管青相贖☆	朐	
p. 1426	☆水蓄其墓	欒	
p. 1426	故使☆水見之	欒	

頁碼	檔案原文摘錄	原書字形	備註
p. 1514	☆天之山	扌 厶 昏	
p. 1577	☆亂三	言 厶 丩	
p. 1651	冬不用☆	夂 厶 嬰	
p. 1651	愛☆也	夂 厶 嬰	
p. 1689	其狀☆然不儻	艮	
p. 1755	必厚其☆	韌	
p. 1775	為其唯厚而及；☆	缶 厶 食	
p. 1775	者☆之	艸 厶 紆	
p. 1775	澤其☆而後之	韌	
p. 1755	見風則☆	彳 厶 厥	
p. 1781	☆辟米	广 厶 走	
p. 1782	薄☆而赤色	禾 厶 羔	