

《呂氏春秋》(標記)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 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》(標記)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》(標記)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，如：人名、地名等特殊情形。例：造字編號 982 的「胷」字，是「胸」的異體字，以標準字「胸」取代。
2. 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》(標記)轉碼分析表

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-----|----------|
| 827 | 絃 | 3 | 8E69 | 9B8C | 絃 | | |
| 957 | 遨 | 3 | 8F4E | 906C | 遨 | | |
| 972 | 犇 | 1 | 8F5D | 7287 | 犇 | | |
| 982 | 胷 | 4 | 8F67 | 80F7 | 胷 | | 胸，異體字問題。 |
| 1024 | 虵 | 1 | 8FB3 | 8675 | 虵 | | |
| 1053 | 傲 | 1 | 8FD0 | 6160 | 傲 | | |
| 1062 | | 4 | 8FD9 | | | 馬𠩺亞 | 駟，異體字問題。 |
| 1099 | 睥 | 5 | 8FFE | 66A4 | 睥 | | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-----|----|
| 1133 | 瓌 | 1 | 9061 | 74CC | 瓌 | | |
| 1243 | 桐 | 4 | 90F1 | 79F1 | 桐 | | |
| 1843 | 擊 | 1 | 94D5 | 22BF2 | | 𠄎𠄎手 | |
| 1857 | 𧈧 | 1 | 94E3 | 86A0 | 𧈧 | | |
| 1885 | 𧈩 | 2 | 9540 | 882D | 𧈩 | | |
| 1937 | 詢 | 2 | 9574 | 8A3D | 詢 | | |
| 2053 | 躡 | 1 | 964B | 8EA7 | 躡 | | |
| 2146 | 𧈭 | 1 | 96CA | 8739 | 𧈭 | | |
| 2162 | 𧈮 | 1 | 96DA | 918E | 𧈮 | | |
| 2610 | 𧈰 | 2 | 99C3 | 9C53 | 𧈰 | | |
| 2612 | 暗 | 1 | 99C5 | 3B10 | | 日𠄎晉 | |
| 3770 | 个 | 8 | 8141 | 4E2A | 个 | | |
| 3780 | 儁 | 2 | 814B | 5101 | 儁 | | |
| 3801 | 𧈱 | 3 | 8160 | 5257 | 𧈱 | | |
| 3804 | 効 | 2 | 8163 | 52B9 | 効 | | |
| 3814 | 却 | 10 | 816D | 5374 | 却 | | |
| 3835 | 噍 | 1 | 81A4 | 564D | 噍 | | |
| 3852 | 媼 | 1 | 81B5 | 5A63 | 媼 | | |
| 3855 | 擘 | 6 | 81B8 | 5B7C | 擘 | | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-------|----------|
| 3873 | 廩 | 1 | 81CA | | | 广 㒼 設 | 廩，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 | 獫 | 2 | 8276 | 7328 | 獫 | | |
| 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 | 縹 | 1 | 82F4 | 7DE5 | 縹 | | |
| 4104 | 莅 | 1 | 8355 | 8385 | 莅 | | |
| 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。 |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-------|----------|
| | | | | | | | 題。 |
| 4164 | 溱 | 1 | 83B3 | 23F6C | | 氵 厶 寒 | |
| 4165 | 獾 | 2 | 83B4 | 8C9B | 獾 | | |
| 4172 | 麇 | 6 | 83BB | 8E77 | 麇 | | |
| 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 | 唵 | 1 | 85EA | 552B | 唵 | | |
| 4629 | 坳 | 2 | 86AD | 577F | 坳 | | |
| 4768 | 妬 | 2 | 8779 | 59AC | 妬 | | 妒，異體字問題。 |
| 4822 | 尫 | 2 | 87D1 | 5C29 | 尫 | | |
| 5021 | 箴 | 1 | 88FB | 25D27 | | 竹 厶 虞 | |
| 5057 | 掇 | 1 | 8960 | 39FB | | 扌 厶 豕 | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-------|-----------------|
| 5091 | 鄒 | 1 | 89A4 | 92E3 | 鄒 | | |
| 5093 | 鳩 | 1 | 89A6 | 9D2A | 鳩 | | 鳩，異體字問題。 |
| 5108 | 惛 | 1 | 89B5 | 60FD | 惛 | | |
| 5302 | 昏 | 2 | 8ADA | 662C | 昏 | | |
| 5306 | | 3 | 8ADE | | | 𠂇 𠂇 閒 | 簡，異體字問題。 |
| 5364 | 𠂇 | 1 | 8B59 | 239D1 | | 𠂇 𠂇 凶 | |
| 5371 | 𠂇 | 3 | 8B60 | 808E | 𠂇 | | |
| 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 | 塾 | 1 | 8C52 | 25095 | | 執 𠂇 皿 | |
| 5578 | 弛 | 1 | 8CB4 | 38AE | | 弓 𠂇 屯 | 弛，異體字問題。 |
| 5730 | 隄 | 3 | 8DAF | 9638 | 隄 | | |
| 5754 | 躡 | 1 | 8DC7 | 2818C | | 足 𠂇 煩 | |
| 5762 | 截 | 1 | 8DCF | 229F5 | | 截 | |

二、手動取代表說明：

《呂氏春秋》(標記)檔案中需手動修改的情況分為以下幾種：

(一)《呂氏春秋》(標記)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「☆」字形手動取代表，請參考附件二。欄位說明如下：

1. 標記行碼：語言所標記檔裡設定的行碼。
2. 檔案原文：摘錄《呂氏春秋》(標記)檔案有「☆」字形的文句。
3. 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》(標記)「☆」字形手動取代表

| 標記行碼 | 檔案原文摘錄 | 原書字形 | 備註 |
|------|---------------------|-------|-------|
| 488 | 腹☆(NB1)[+prop] | 黃厶享 | 有 3 處 |
| 968 | 禽滑☆(NB1)[+prop] | 毆厶康 | 有 2 處 |
| 1214 | ☆(U) 為(VG) 盲(NI) | 目厶蔑 | |
| 1219 | 多(VH1) ☆(VH1)[+p1] | 允厶重 | 改為𠂔 |
| 2599 | 麋☆(NA2)[+attr] | 革厶各 | |
| 2685 | 化(VH1) 為(VG) ☆(NA1) | 𧈧 | |
| 3208 | 頡☆(VH1)[+co] | 吾厶午 | |
| 3558 | 胎☆(NA1)[+co] | 月厶賣 | |
| 4300 | ☆(VP)[+nv, +jj] | 磨厶𠂔 | |
| 4450 | 過(VH1N)[+nv] ☆(NA1) | 册厶頁 | |
| 4644 | ☆研(VA)[+rd] | 磨厶𠂔 | |
| 4970 | 釜☆(VA)[+co, +nv] | 鬲厶瓦 | |
| 5910 | 結(VP) ☆ | 冈厶不 | 改為罍 |
| 5910 | ☆(NA2)[+attr] | 冈厶亡 | 改為𠂔厶亡 |
| 6052 | ☆(NB4)[+prop] 也(T) | 革厶處 | |
| 6534 | 角☆(NA2)[+attr] | 角厶妻厶才 | |
| 7334 | 不(DC) ☆ | 月厶侯 | |

| 標記行碼 | 檔案原文摘錄 | 原書字形 | 備註 |
|-------|--------------------------------|------|-------|
| 7346 | 洞庭(NB4)[+prop] 之(T) ☆ (NA1) | 魚△專 | |
| 7374 | 陽☆(NB4)[+prop] | 櫟 | |
| 9837 | ☆(U) | 詭 | |
| 10994 | 多☆(NB2)[+prop] | 嬰△頁 | |
| 11928 | 蹠☆(NA2)[+attr] | 疒△稭 | |
| 12091 | 口(NA1) ☆(VH1) | 口△昏 | |
| 12651 | ☆(NB4)[+prop] | 𠂔穴忄分 | |
| 12682 | 絳(NB4)[+prop] 、 ☆(NB4) | 𠂔穴忄分 | |
| 13032 | 丹績(NA2)[+attr] 之(T) ☆ | 衤△旬 | |
| 13701 | ☆為(NI)[+attr] | 木△異 | |
| 15200 | 相(VC1) ☆(NA1) | 月△尻 | |
| 15202 | 贖☆(NA1)[+co] | 胸 | |
| 15269 | ☆水(NB4)[+prop] | 樂 | 有 2 處 |
| 16223 | ☆天(NB4)[+prop] | 扌△昏 | |
| 16929 | ☆(VE)[+jj] 亂(NI) 三 | 言△丩 | |
| 17833 | 不(DC) 用(VC1) ☆(NA2) | 𠂔△嬰 | |
| 17834 | 愛(VH1N) ☆(NA2) 也(T) | 𠂔△嬰 | |
| 18242 | ☆然(VI)[+poly] | 眼 | |
| 18664 | 其(NH) ☆(NA2) | 鞞 | |
| 18666 | ☆(VH1)[+jj] 者(NH) | 缶△食 | |
| 18666 | ☆(VC1) 之(NH) | 艹△紕 | |
| 18668 | 澤(VP)[+nv] 其(NH) ☆ | 鞞 | |
| 18695 | 見(VK) 風(NA2) 則(C) ☆ (VA) | 彳△厥 | |
| 18832 | ☆(U) 辟米 | 疒△走 | |

| 標記行碼 | 檔案原文摘錄 | 原書字形 | 備註 |
|-------|----------------|------|----|
| 18864 | 薄(VH1N) ☆(NA2) | 禾𪔐羔 | |