

《韓非子》舊版造字轉碼說明

中研院資訊所文獻處理實驗室
中研院史語所漢籍電子文獻工作小組
2008/6/10 陳維君 製作

一、《韓非子》一書(韓非子.xml)使用舊版造字 71 個，字頻 243 次，詳如附件一。這 71 個造字中，59 個可轉成 Windows XP 能顯示的字，字頻 218 次；另外 12 個字必須轉成構字式，字頻 25 次。

二、附件一的造字分析表說明如下：

甲、編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、造字：舊版造字

丙、字頻(txt)：造字在「.txt」文件的出現次數

丁、字頻(xml)：造字在「.xml」文件的出現次數

戊、Big5：造字的 Big5 碼

己、Unicode：造字所對應的 Unicode 碼

庚、WinXP：造字在 Windows XP 的對應字形

辛、構字式：Windows XP 若無對應字形，則改採用構字式

壬、備註凡例：

- 1、校對問題，舊版漢籍錯字，可用程式全部取代：在舊版漢籍電子文獻中即存在的錯字，因校對時的疏漏而未更正，持續留存在新版漢籍電子文獻中；若該造字的所有頻次，皆屬於錯誤使用的錯字情形，可以用程式全部取代為正確字形。如編號 2013 的「𪛗」字，原字為「𪛗」。
- 2、校對問題，舊版漢籍錯字，必須人工取代：在舊版漢籍電子文獻中即存在的錯字，因校對時的疏漏而未更正，持續留存在新版漢籍電子文獻中，然而在所有頻次中，只有一部分是錯字，另一部分是正確的，因此無法全部取代，必須人工一一比對，將錯誤的部分更正，如編號 4190「郟」，共出現 6 次，但原書字形有 1 次為「衛」字，5 次為「郟」字，必須手動取代。
- 3、異體字問題：新版漢籍考量到使用者檢索及使用時的便利性，將用字原則改為除專詞等特殊情形之外，一律改用標準字呈現。如編號 982 的「𪛗」係「胸」字之異體，故以「胸」字取代。
- 4、Unicode 字型呈現差異：Unicode 字型與舊漢籍造字有些微差異，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 4242 的「鬪」字，

Unicode 字型呈現為「鬪」，實際上仍為同一字。

- 三、 Unicode 目前收錄的漢字總數為 70194，分屬於三個不同區段，詳如表一。目前 Windows XP 只支援 CJK 認同表意文字區的 20902 個字，內碼為 4E00-9FFF。所以造字編號 790 的「𠄎」字，Unicode 編碼為 4039，由於 Windows XP 並不支援，仍須使用構字式「目𠄎夾」。

表一、Unicode 的字數及編碼區段

| Unicode | 字集子集合 | 新增字數 | 新增編碼區段 | 總字數 | WinXP |
|---------|------------------|-------|-------------|-------|-------|
| 1.1 版 | CJK 認同表意文字區 | 20902 | 4E00-9FFF | 20902 | 支援 |
| 3.0 版 | CJK 認同表意文字擴充 A 區 | 6582 | 3400-4DFF | 27484 | 不支援 |
| 3.1 版 | CJK 認同表意文字擴充 B 區 | 42710 | 20000-2A6D6 | 70194 | 不支援 |

- 四、 txt 檔與 xml 檔頻次不同且造字分析表「備註」欄有備註問題情況者，詳細取代內容可參考附件二的手動取代表。附件二的手動取代表的欄位說明如下：

- 甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
- 乙、 造字：舊版造字字形。
- 丙、 取代：當實際使用字形與造字不同時，在此標出。
- 丁、 檔案原文摘錄：摘錄檔案(韓非子.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。
- 戊、 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。

- 五、 《韓非子》一書(韓非子.xml)中未確定字形「●」共出現 22 次，詳細取代內容可參考附件三未確定字形「●」手動取代表。附件三未確定字形「●」手動取代表的欄位說明如下：

- 甲、 頁碼：「●」字形所在位置之原書頁碼。
- 乙、 檔案原文摘錄：摘錄檔案(韓非子.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。
- 丙、 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。

附件一、《韓非子》造字分析表

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXp | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-------|--------------------------|
| 790 | 映 | 3 | 3 | 8E44 | 4039 | | 目 𠃉 夾 | 部分為犁，校對問題，舊版漢籍錯字，必須人工取代。 |
| 826 | 癱 | 1 | 1 | 8E68 | 7655 | 癱 | | |
| 982 | 胷 | 1 | 1 | 8F67 | 80F7 | 胷 | | 胸，異體字問題。 |
| 1040 | 聃 | 1 | 1 | 8FC3 | 803C | 聃 | | 聃，異體字問題。 |
| 1044 | 儻 | 1 | 1 | 8FC7 | 511B | 儻 | | |
| 1048 | 頷 | 1 | 1 | 8FCB | 981F | 頷 | | |
| 1053 | 傲 | 2 | 2 | 8FD0 | 6160 | 傲 | | |
| 1067 | 碁 | 1 | 1 | 8FDE | 68CA | 碁 | | |
| 1097 | 嶠 | 1 | 1 | 8FFC | 37ED | | 山 𠃉 昏 | |
| 1394 | 啗 | 3 | 3 | 91EB | 35D6 | | 口 𠃉 啗 | |
| 1583 | 談 | 1 | 1 | 934C | 46DF | | 言 𠃉 夾 | |
| 1912 | 襍 | 1 | 1 | 955B | 894D | 襍 | | |
| 1981 | 讐 | 6 | 6 | 95C2 | 8B90 | 讐 | | 讐，異體字問題。 |
| 1982 | 調 | 1 | 1 | 95C3 | 8B87 | 調 | | |
| 2013 | 𡗗 | 1 | 1 | 95E2 | 291F5 | | 非 𠃉 古 | 讐，校對問題，舊版漢籍錯字，可用程式全部取代。 |
| 2811 | | 1 | 1 | 9AEF | | | 𠃉 𠃉 田 | 雷，異體字問題。 |

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXp | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-------|-----------------------------|
| 2817 | 窻 | 1 | 1 | 9AF5 | 3AC4 | | 窻 | |
| 3791 | 況 | 2 | 2 | 8156 | 51B5 | 況 | | 況，異體字問題。 |
| 3806 | 勳 | 1 | 1 | 8165 | 52E0 | 勳 | | |
| 3830 | 咤 | 1 | 1 | 817D | 54A4 | 咤 | | |
| 3846 | 壻 | 1 | 1 | 81AF | 58FB | 壻 | | Unicode 呈現差異。 |
| 3855 | 擘 | 2 | 2 | 81B8 | 5B7C | 擘 | | |
| 3868 | 卮 | 11 | 11 | 81C5 | 5DF5 | 卮 | | |
| 3873 | 廩 | 2 | 2 | 81CA | | | 广 㒼 𠂔 | 廩，異體字問題。 |
| 3883 | 徧 | 7 | 7 | 81D4 | 5FA7 | 徧 | | |
| 3929 | 黎 | 4 | 4 | 8243 | 68C3 | 黎 | | |
| 3930 | 攄 | 2 | 2 | 8244 | 6901 | 攄 | | |
| 3935 | 槩 | 1 | 1 | 8249 | 69E9 | 槩 | | |
| 3937 | 檝 | 2 | 2 | 824B | 6A9D | 檝 | | |
| 3966 | 烟 | 1 | 1 | 8268 | 70DF | 烟 | | |
| 3976 | 犁 | 1 | 1 | 8272 | 7282 | 犁 | | 目 㒼 夾，校對問題，舊版漢籍錯字，可用程式全部取代。 |
| 3977 | 狗 | 1 | 1 | 8273 | 72E5 | 狗 | | |
| 3996 | 踈 | 4 | 4 | 82A8 | 758E | 踈 | | |

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXp | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-----|-------------------------|
| 4035 | 窳 | 2 | 2 | 82CF | 7A7D | 窳 | | |
| 4038 | 竈 | 10 | 10 | 82D2 | 7AC8 | 竈 | | |
| 4051 | 籛 | 1 | 1 | 82DF | 7C35 | 籛 | | |
| 4070 | 縹 | 1 | 1 | 82F2 | 7DE4 | 縹 | | |
| 4104 | 蒞 | 7 | 7 | 8355 | 8385 | 蒞 | | 蒞，異體字問題。 |
| 4131 | 劬 | 2 | 2 | 8370 | 8842 | 劬 | | |
| 4132 | | 1 | 1 | 8371 | | | 血𠂇𠂇 | 羈，校對問題，舊版漢籍錯字，可用程式全部取代。 |
| 4134 | 衛 | 46 | 46 | 8373 | 885E | 衛 | | |
| 4136 | 衰 | 3 | 3 | 8375 | 887A | 衰 | | |
| 4139 | 袴 | 11 | 11 | 8378 | 88B4 | 袴 | | |
| 4140 | 衽 | 1 | 1 | 8379 | 88B5 | 衽 | | |
| 4145 | 羈 | 1 | 1 | 837E | 898A | 羈 | | Unicode 呈現差異。 |
| 4172 | 蹶 | 2 | 2 | 83BB | 8E77 | 蹶 | | 蹶，異體字問題。 |
| 4184 | 迹 | 9 | 9 | 83C7 | 8FF9 | 迹 | | |
| 4190 | 邳 | 6 | 6 | 83CD | 90C4 | 邳 | | 部分為衛，校對問題，舊版漢籍錯字，必須人工取代 |
| 4216 | 鞮 | 1 | 1 | 83E7 | 97C8 | 鞮 | | |
| 4226 | 凜 | 5 | 5 | 83F1 | 98E1 | 凜 | | |

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXp | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-------|---------------|
| 4242 | 鬪 | 13 | 13 | 8442 | 9B2D | 鬪 | | Unicode 呈現差異。 |
| 4243 | 黼 | 2 | 2 | 8443 | 9B34 | 黼 | | |
| 4265 | 鼃 | 4 | 4 | 8459 | 9F03 | 鼃 | | |
| 4267 | 鼈 | 3 | 3 | 845B | 9F08 | 鼈 | | |
| 4501 | 獫 | 1 | 1 | 85CA | 7356 | 獫 | | |
| 4533 | 唸 | 1 | 1 | 85EA | 552B | 唸 | | |
| 4768 | 妬 | 6 | 6 | 8779 | 59AC | 妬 | | 妒，異體字問題。 |
| 4904 | 尅 | 2 | 2 | 8864 | 5C05 | 尅 | | 剋，異體字問題。 |
| 4996 | 斲 | 3 | 3 | 88E2 | 65B5 | 斲 | | |
| 5034 | 弃 | 3 | 3 | 8949 | 5F03 | 弃 | | 棄，異體字問題。 |
| 5091 | 鄒 | 1 | 1 | 89A4 | 92E3 | 鄒 | | |
| 5108 | 惛 | 5 | 5 | 89B5 | 60FD | 惛 | | |
| 5300 | 鬪 | 1 | 1 | 8AD8 | 9B2A | 鬪 | | 鬪，異體字問題。 |
| 5302 | 昏 | 2 | 2 | 8ADA | 662C | 昏 | | 昏，異體字問題。 |
| 5396 | 鎔 | 1 | 1 | 8B79 | 938B | 鎔 | | |
| 5578 | 弛 | 5 | 5 | 8CB4 | 38AE | | 弓 厶 屯 | 弛，異體字問題。 |
| 5705 | 廐 | 3 | 3 | 8D74 | 5ED0 | 廐 | | 廐，異體字問題。 |
| 5709 | 蟀 | 1 | 1 | 8D78 | 872F | 蟀 | | |

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXp | 構字式 | 備註 |
|------|----|----------|----------|------|---------|-------|-----|----|
| 5722 | 朥 | 5 | 5 | 8DA7 | | 朥 | | |
| 5726 | 呂 | 1 | 1 | 8DAB | | | 呂 | |
| 5752 | 摠 | 1 | 1 | 8DC5 | 63D4 | 摠 | | |

附件二、《韓非子》手動取代表

| 編號 | 造字 | 取代 | 檔案原文摘錄 | 手動取代結果 |
|------|-------|--------|-----------------|------------------------|
| 790 | 目 㒼 夾 | | 則__其一目 ，兩目__ | 則目 㒼 夾其一目 ，兩目 目 㒼 夾 |
| | | 犁 | 或宦於荊，__鉏曰： | 或宦於荊，犁鉏曰： |
| 4190 | 郄 | 衛 | 郄州吁殺其君完 | 衛州吁殺其君完 |
| | | | 費無忌教郄宛而令尹 | 費無忌教郄宛而令尹 |
| | | | 郄宛新事令尹 | 郄宛新事令尹 |
| | | | 因令之為具於郄宛之家 | 因令之為具於郄宛之家 |
| | | | 舉兵而誅郄宛 | 舉兵而誅郄宛 |
| | | 則交必有郄矣 | 則交必有郄矣 | |

附件三、《韓非子》「●」字形手動取代表

| 頁碼 | 檔案原文摘錄 | 手動取代結果 |
|-----|----------|--------------|
| 37 | 虛處則●然 | 虛處則卣 㒼 亥然 |
| 124 | 其鬪●● | 其鬪口 㒼 顏口 㒼 顏 |
| 298 | 而●小變而失長便 | 而車 㒼 辰小變而失長便 |
| 387 | 漆其首以為●器 | 漆其首以為 㒼 㒼 器 |
| 594 | 有燒倉廡●者 | 有燒倉廡穴 㒼 舛者 |
| 645 | ●擻而置之 | 扌 㒼 騫擻而置之 |
| 670 | ●危坐子臯 | 𠂔危坐子臯 |
| | 故昭卯五乘而履● | 故昭卯五乘而履尸 㒼 橋 |
| 673 | 與●危子榮衣 | 與𠂔危子榮衣 |

| 頁碼 | 檔案原文摘錄 | 手動取代結果 |
|------|-----------|--------------------------|
| 677 | 所●者守門 | 所 明 者守門 |
| | ●危引之而逃 | 明 危引之而逃 |
| | 子臬問●危曰 | 子臬問 明 危曰 |
| | 而親●子之足 | 而親 明 子之足 |
| | ●危曰：『吾斷足也 | 明 危曰：『吾斷足也 |
| 691 | 吾將何●以履之 | 吾將何 尸 △橋以履之 |
| 762 | 成竅●下 | 成竅 撤 下 |
| 775 | 人主之所以自淺●者 | 人主之所以自淺 娟 者 |
| 844 | 吾聞龐●氏之子不孝 | 吾聞龐 米 △間氏之子不孝 |
| | 問龐●氏子 | 問龐 米 △間氏子 |
| 932 | ●尊嬖臣而匹上卿 | 無 尊嬖臣而匹上卿 |
| 1103 | 不●瘞則寢益 | 不 才 △副瘞則寢益 |