

# 《荀子》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  
2010/04/13 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 1339 的「儻」字，Unicode 編碼是 34A9，由於 XP 並不支援，仍需使用構字式「亻 𠂔 蠡」。

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：

《荀子》使用舊版造字 72 個，字頻 209 次，這 72 個造字中，66 個可轉成 Windows XP 能顯示的字，字頻 193 次；另外 6 個字必須轉成構字式，字頻 16 次。

轉碼完成製作轉碼分析表，請參考附件一《荀子》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 5299 的「𣪠」字，是「𣪠」的異體字，以標準字「𣪠」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 1836 的「口𣪠丹」字，而原書使用之字為「口𣪠丹」，查詢教育部異體字字典確認「口𣪠丹」並非「口𣪠丹」之異體字，因此歸為錯字，以程式取代為正確字形「口𣪠丹」。
  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 3978 的「猪」字，Unicode 字型呈現為「猪」，實際上仍為同一字。

附件一、《荀子》轉碼分析表

| 編號   | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-----|----|
| 813  | 𣪠  | 1  | 8E5B | 65E4    | 𣪠     |     |    |
| 972  | 犇  | 5  | 8F5D | 7287    | 犇     |     |    |
| 1039 | 𣪠  | 1  | 8FC2 | 69C0    | 𣪠     |     |    |
| 1059 | 𣪠  | 1  | 8FD6 | 51A3    | 𣪠     |     |    |

| 編號   | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式   | 備註                |
|------|----|----|------|---------|-------|-------|-------------------|
| 1077 | 稟  | 1  | 8FE8 | 7A3E    | 稟     |       |                   |
| 1339 | 儼  | 2  | 91B4 | 34A9    |       | 亻 厶 𧈧 |                   |
| 1565 | 繇  | 2  | 92F9 | 7DDC    | 繇     |       |                   |
| 1836 | 𠵼  | 2  | 94CE |         |       | 口 厶 丹 | 口 厶 丹，錯字，以程式全部取代。 |
| 1885 | 𧈧  | 4  | 9540 | 882D    | 𧈧     |       |                   |
| 1937 | 詢  | 1  | 9574 | 8A3D    | 詢     |       |                   |
| 1981 | 讐  | 3  | 95C2 | 8B90    | 讐     |       |                   |
| 2146 | 蝸  | 1  | 96CA | 8739    | 蝸     |       |                   |
| 3039 | 覩  | 1  | 9C77 | 89A9    | 覩     |       |                   |
| 3061 | 僂  | 1  | 9CAF | 5044    | 僂     |       |                   |
| 3799 | 刼  | 1  | 815E | 5226    | 刼     |       |                   |
| 3831 | 喆  | 1  | 817E | 5586    | 喆     |       |                   |
| 3835 | 噍  | 2  | 81A4 | 564D    | 噍     |       |                   |
| 3855 | 擘  | 2  | 81B8 | 5B7C    | 擘     |       |                   |
| 3883 | 徧  | 10 | 81D4 | 5FA7    | 徧     |       |                   |
| 3888 | 愆  | 1  | 81D9 | 60B3    | 愆     |       |                   |

| 編號   | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註                  |
|------|----|----|------|---------|-------|-----|---------------------|
| 3924 | 朥  | 1  | 81FD | 671E    | 朥     |     |                     |
| 3930 | 捰  | 4  | 8244 | 6901    | 捰     |     |                     |
| 3935 | 槩  | 2  | 8249 | 3BA3    | 槩     |     | Unicode 字型<br>呈現差異。 |
| 3937 | 檝  | 1  | 824B | 6A9D    | 檝     |     |                     |
| 3951 | 湏  | 1  | 8259 | 6E4F    | 湏     |     |                     |
| 3974 | 牀  | 6  | 8270 | 7240    | 牀     |     |                     |
| 3978 | 猪  | 2  | 8274 | 732A    | 猪     |     | Unicode 字型<br>呈現差異。 |
| 3981 | 珉  | 3  | 8277 | 73C9    | 珉     |     |                     |
| 4014 | 蓋  | 1  | 82BA | 8462    | 蓋     |     |                     |
| 4040 | 竝  | 12 | 82D4 | 7ADD    | 竝     |     |                     |
| 4103 | 苙  | 1  | 8354 | 831D    | 苙     |     |                     |
| 4104 | 苘  | 1  | 8355 | 8385    | 苘     |     |                     |
| 4124 | 蠱  | 3  | 8369 | 8771    | 蠱     |     |                     |
| 4146 | 覩  | 2  | 83A1 | 89A9    | 覩     |     |                     |
| 4156 | 諂  | 1  | 83AB | 8B1F    | 諂     |     |                     |
| 4172 | 屨  | 4  | 83BB | 8E77    | 屨     |     |                     |

| 編號   | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式   | 備註       |
|------|----|----|------|---------|-------|-------|----------|
| 4184 | 迹  | 4  | 83C7 | 8FF9    | 迹     |       |          |
| 4190 | 邳  | 1  | 83CD | 90C4    | 邳     |       |          |
| 4222 | 頰  | 1  | 83ED | 9847    | 頰     |       |          |
| 4226 | 凜  | 1  | 83F1 | 98E1    | 凜     |       |          |
| 4250 | 鱈  | 7  | 844A | 9C0C    | 鱈     |       |          |
| 4256 | 鷄  | 3  | 8450 | 9DC4    | 鷄     |       |          |
| 4265 | 鼃  | 1  | 8459 | 9F03    | 鼃     |       |          |
| 4267 | 鼈  | 8  | 845B | 9F08    | 鼈     |       |          |
| 4287 | 織  | 1  | 846F | 7E8E    | 織     |       |          |
| 4312 | 皞  | 1  | 84AA | 66AD    | 皞     |       | 皞，異體字問題。 |
| 4342 | 倮  | 1  | 84C8 | 502E    | 倮     |       |          |
| 4458 | 勸  | 4  | 857D | 52CC    | 勸     |       |          |
| 4768 | 妬  | 8  | 8779 | 59AC    | 妬     |       |          |
| 4855 | 汩  | 1  | 87F2 | 3CC5    |       | 汩 𠄎 不 |          |
| 4996 | 斲  | 2  | 88E2 | 65B5    | 斲     |       |          |
| 5034 | 弃  | 8  | 8949 | 5F03    | 弃     |       | 棄，異體字問題。 |

| 編號   | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式   | 備註              |
|------|----|----|------|---------|-------|-------|-----------------|
| 5133 | 憇  | 1  | 89CE | 6199    | 憇     |       |                 |
| 5169 | 戇  | 2  | 89F2 | 6239    | 戇     |       |                 |
| 5289 | 慤  | 3  | 8ACD | 6164    | 慤     |       |                 |
| 5299 | 戟  | 2  | 8AD7 | 39B8    |       | 卓 厶 戈 | 戟，異體字問題。        |
| 5300 | 鬪  | 27 | 8AD8 | 9B2A    | 鬪     |       |                 |
| 5302 | 昏  | 1  | 8ADA | 662C    | 昏     |       |                 |
| 5351 | 肫  | 1  | 8B4C | 80AC    | 肫     |       |                 |
| 5355 | 𡗗  | 1  | 8B50 | 20AE4   | 𡗗     |       |                 |
| 5434 | 湟  | 1  | 8BC1 | 6E7C    | 湟     |       |                 |
| 5655 | 𪗇  | 1  | 8D42 | 56BB    | 𪗇     |       | 𪗇，異體字問題。        |
| 5687 | 𪗈  | 9  | 8D62 | 4AC9    |       | 𪗈     |                 |
| 5689 | 𪗉  | 1  | 8D64 | 20087   |       | 𪗉     |                 |
| 5701 | 徧  | 8  | 8D70 | 5FA7    | 徧     |       |                 |
| 5705 | 廐  | 1  | 8D74 | 5ED0    | 廐     |       | Unicode 字型呈現差異。 |
| 5706 | 敍  | 2  | 8D75 | 654D    | 敍     |       | 敍，異體字問題。        |

| 編號   | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-----|----|
| 5717 | 𨺗  | 1  | 8DA2 | 967F    | 𨺗     |     |    |
| 5722 | 𨺗  | 1  | 8DA7 | 671E    | 𨺗     |     |    |
| 5730 | 𨺗  | 4  | 8DAF | 9638    | 𨺗     |     |    |
| 5749 | 𨺗  | 1  | 8DC2 | 7257    | 𨺗     |     |    |
| 5752 | 𨺗  | 1  | 8DC5 | 63D4    | 𨺗     |     |    |

## 二、手動取代表說明：

《荀子》檔案中出現「☆」字形，共計 87 字，比對原書後以 Unicod 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《荀子》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。

### 附件二、《荀子》「☆」字手動取代表

| 頁碼    | 檔案原文摘錄 | 原書字形 | 備註 |
|-------|--------|------|----|
| p. 2  | ☆、水為之  | 氷    |    |
| p. 2  | 雖有槁☆   | 暴    |    |
| p. 23 | ☆☆些些   | 喻喻   |    |
| p. 42 | 而不☆    | 暴    |    |

| 頁碼     | 檔案原文摘錄  | 原書字形  | 備註    |
|--------|---------|-------|-------|
| p. 43  | 非驕☆也    | 暴     |       |
| p. 44  | 大心則慢而☆  | 暴     |       |
| p. 57  | 乳☆不遠遊   | 彳 厶 苟 |       |
| p. 57  | ☆毳      | 彳 厶 苟 | 有 3 處 |
| p. 58  | ☆☆然惟利   | 忤忤    |       |
| p. 59  | 儵☆者     | 魚 厶 本 |       |
| p. 60  | 僑☆      | 暴     | 有 2 處 |
| p. 68  | 方知畜鷄☆   | 彳 厶 苟 |       |
| p. 89  | ☆唯則節    | 瞻     |       |
| p. 108 | 而忍☆詢    | 言 厶 隹 |       |
| p. 116 | ☆國      | 暴     | 有 9 處 |
| p. 120 | 可炊而☆也   | 僥     |       |
| p. 126 | 而非☆也    | 暴     |       |
| p. 136 | ☆回天下於   | 口 厶 冂 |       |
| p. 141 | ☆☆兮其能長久 | 馱馱    |       |
| p. 147 | 疆☆      | 暴     | 有 2 處 |
| p. 149 | ☆然若     | 彳 厶 患 |       |
| p. 149 | 無所擬☆    | 作 厶 心 |       |
| p. 171 | 禁☆      | 暴     | 有 5 處 |
| p. 190 | 可以誅☆    | 暴     |       |
| p. 190 | 日日☆露    | 暴     |       |



| 頁碼     | 檔案原文摘錄 | 原書字形  | 備註    |
|--------|--------|-------|-------|
| p. 211 | ☆☆如丘山  | 暴暴    |       |
| p. 225 | 強☆之國   | 暴     | 有 5 處 |
| p. 225 | 足以☆炙之  | 暴     |       |
| p. 257 | 賤之如☆   | 彳 厶 匡 |       |
| p. 281 | ☆然而齒墮  | 齟     |       |
| p. 294 | 事☆君    | 暴     | 有 2 處 |
| p. 298 | 不敢☆虎   | 暴     |       |
| p. 316 | 冠☆帶劍   | 革 厶 由 |       |
| p. 331 | 宛鉅鐵☆   | 鉞     |       |
| p. 335 | ☆悍勇力   | 暴     |       |
| p. 342 | ☆察之威   | 暴     | 有 3 處 |
| p. 393 | ☆屨，殺   | 彳 厶 封 |       |
| p. 393 | 惠☆而寬賊  | 暴     |       |
| p. 393 | 夫征☆誅悍  | 暴     |       |
| p. 405 | 故盜不☆   | 窃     | 改為竊   |
| p. 411 | ☆貪利    | 暴     |       |
| p. 411 | 舌☆，是辱  | 彳 厶 舉 |       |
| p. 428 | ☆慢恣睢   | 暴     |       |
| p. 438 | ☆笑、哭泣  | 讖     |       |
| p. 459 | 怒而☆亂畏之 | 暴     |       |
| p. 460 | 帶甲嬰☆   | 革 厶 由 |       |

| 頁碼     | 檔案原文摘錄 | 原書字形       | 備註 |
|--------|--------|------------|----|
| p. 464 | ☆杙、拊臬  | 鞞          |    |
| p. 464 | 積意☆☆乎  | 言厶犀<br>言厶犀 |    |
| p. 534 | 〔尚〕☆机  | 尙          |    |
| p. 538 | 而歸於☆   | 暴          |    |
| p. 547 | 眾者☆寡   | 暴          |    |
| p. 558 | 不得排☆   | 木厶敬        |    |
| p. 558 | 莊君之☆   | 忽          |    |
| p. 590 | 充盈乎大☆  | 宀厶禹        |    |
| p. 613 | ☆泮殺內   | 氷          |    |
| p. 629 | ☆如也    | 嶼          |    |
| p. 647 | 藜羹不☆   | 糲          |    |