

《歧路燈》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
2010/02/04 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 1691 的「鬚」字，Unicode 編碼是 4BFC，由於 XP 並不支援，仍需使用構字式「鬚𠄎狄」。

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：

《歧路燈》使用舊版造字 139 個，字頻 3015 次，這 139 個造字中，127 個可轉成 Windows XP 能顯示的字，字頻 2975 次；另外 12 個字必須轉成構字式，字頻 40 次。

轉碼完成製作轉碼分析表，請參考附件一《歧路燈》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. 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。
3. 待造字：Unicode 及漢字構形資料庫皆未收錄的舊漢籍造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 3893 的「彳𠂔妄」。


附件一、《歧路燈》轉碼分析表

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-----|----------|
| 1153 | 坳 | 1 | 9075 | 362D | | 土𠂔幻 | |
| 1201 | 昵 | 1 | 90C7 | 7724 | 昵 | | |
| 1278 | 碍 | 7 | 9155 | 788D | 碍 | | |
| 1634 | 忪 | 7 | 93A1 | 5FEC | 忪 | | |
| 1641 | 胆 | 11 | 93A8 | 80C6 | 胆 | | 膽，異體字問題。 |
| 1642 | 脉 | 8 | 93A9 | 8109 | 脉 | | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|------------------------------------|----------|
| 1691 | 髡 | 1 | 93DA | 4BFC | | 髡𠂇狄 | |
| 2046 | 踪 | 24 | 9644 | 8E2A | 踪 | | |
| 2206 | 袂 | 1 | 9747 | 8884 | 袂 | | |
| 2361 | 鎗 | 1 | 9845 | 93F3 | 鎗 | | |
| 2374 | 冢 | 2 | 9852 | 51A1 | 冢 | | |
| 2408 | 鬧 | 10 | 9874 | 9599 | 鬧 | | 鬧，異體字問題。 |
| 2415 | 鬪 | 2 | 987B | 95D8 | 鬪 | | 鬪，異體字問題。 |
| 2448 | 隼 | 3 | 98BE | 96BD | 隼 | | |
| 2592 | 贓 | 9 | 99B1 | 8D1C | 贓 | | |
| 2758 | 獻 | 1 | 9ABA | 732E | 獻 | | 獻，異體字問題。 |
| 2812 | 舉 | 4 | 9AF0 | 64E7 | 舉 | | |
| 2883 | 涂 | 1 | 9B78 | 51C3 | 涂 | | |
| 2997 | 鹽 | 2 | 9C4D | 25081 | | 形土 ^レ 口皿 ^口 | |
| 2999 | 咆 | 5 | 9C4F | 5523 | 咆 | | |
| 3016 | 闕 | 1 | 9C60 | 28D91 | | 門 ^門 臭 | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-----|------|---------|-------|-----|----------|
| 3145 | 擻 | 1 | 9D44 | 64C0 | 擻 | | |
| 3160 | 坟 | 3 | 9D53 | 575F | 坟 | | |
| 3168 | 瘠 | 1 | 9D5B | 24EA5 | | 疒△答 | |
| 3177 | 蟻 | 3 | 9D64 | 7F4E | 蟻 | | |
| 3251 | 灯 | 5 | 9DD0 | 706F | 灯 | | |
| 3788 | 慤 | 1 | 8153 | 22871 | | 殼△心 | |
| 3789 | 冲 | 24 | 8154 | 51B2 | 冲 | | 沖，異體字問題。 |
| 3791 | 况 | 31 | 8156 | 51B5 | 况 | | 況，異體字問題。 |
| 3792 | 凉 | 6 | 8157 | 51C9 | 凉 | | 涼，異體字問題。 |
| 3794 | 凑 | 21 | 8159 | 51D1 | 凑 | | 湊，異體字問題。 |
| 3796 | 决 | 1 | 815B | 51B3 | 决 | | 決，異體字問題。 |
| 3797 | 凭 | 2 | 815C | 51ED | 凭 | | |
| 3805 | 勅 | 2 | 8164 | 52C5 | 勅 | | |
| 3814 | 却 | 701 | 816D | 5374 | 却 | | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-------|----------|
| 3815 | 𡗗 | 1 | 816E | 537A | 𡗗 | | 𡗗，異體字問題。 |
| 3817 | 厠 | 12 | 8170 | 53A0 | 厠 | | |
| 3818 | 厰 | 1 | 8171 | 53AB | 厰 | | |
| 3819 | 廡 | 69 | 8172 | 53AE | 廡 | | 廡，異體字問題。 |
| 3824 | 叙 | 11 | 8177 | 53D9 | 叙 | | |
| 3825 | 疊 | 14 | 8178 | 53E0 | 疊 | | |
| 3829 | 咏 | 3 | 817C | 548F | 咏 | | |
| 3849 | 獎 | 7 | 81B2 | 596C | 獎 | | 獎，異體字問題。 |
| 3851 | 姹 | 1 | 81B4 | 59F9 | 姹 | | |
| 3858 | 寃 | 21 | 81BB | 5BC3 | 寃 | | 寃，異體字問題。 |
| 3883 | 徧 | 1 | 81D4 | 5FA7 | 徧 | | |
| 3890 | 憑 | 2 | 81DB | 6142 | 憑 | | |
| 3893 | | 1 | 81DE | | | 亻 厶 妄 | 待造字。 |
| 3899 | 担 | 3 | 81E4 | 62C5 | 担 | | |
| 3909 | 擡 | 42 | 81EE | 64E1 | 擡 | | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-----|------|---------|-------|---|---------------------|
| 3921 | 皙 | 2 | 81FA | 6673 | 皙 | | |
| 3930 | 捰 | 8 | 8244 | 6901 | 捰 | | |
| 3934 | 楈 | 19 | 8248 | 69C5 | 楈 | | |
| 3938 | 橈 | 8 | 824C | 6AC8 | 橈 | | |
| 3942 | 毡 | 3 | 8250 | 6BE1 | 毡 | | |
| 3966 | 烟 | 40 | 8268 | 70DF | 烟 | | |
| 3975 | 牕 | 2 | 8271 | 7255 | 牕 | | |
| 3976 | 犁 | 3 | 8272 | 7282 | 犁 | | |
| 3978 | 猪 | 34 | 8274 | 732A | 猪 | | Unicode 字型 呈現差異。 |
| 3979 | 猫 | 4 | 8275 | 732B | 猫 | | |
| 4000 | 疴 | 1 | 82AC | 75B4 | 疴 | | |
| 4016 | 着 | 440 | 82BC | 7740 | 着 | | |
| 4028 | 稟 | 166 | 82C8 | 7980 | 稟 | | |
| 4030 | 秆 | 2 | 82CA | 79C6 | 秆 | | |
| 4032 | 秣 | 1 | 82CC | 2578A | | 禾  未 | 秣，異體字問題。 |
| 4036 | 窰 | 4 | 82D0 | 7AB0 | 窰 | | |
| 4038 | 竈 | 1 | 82D2 | 7AC8 | 竈 | | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-----|------|---------|-------|-----|----------|
| 4041 | 豎 | 6 | 82D5 | 7AEA | 豎 | | |
| 4042 | 笋 | 9 | 82D6 | 7B0B | 笋 | | |
| 4043 | 筭 | 1 | 82D7 | 7B53 | 筭 | | |
| 4066 | 綉 | 2 | 82EE | 7D5D | 綉 | | |
| 4067 | 綉 | 29 | 82EF | 7D89 | 綉 | | |
| 4068 | 綫 | 11 | 82F0 | 7DAB | 綫 | | |
| 4073 | 纒 | 2 | 82F5 | 7E6E | 纒 | | |
| 4074 | 縑 | 1 | 82F6 | 7E27 | 縑 | | |
| 4085 | 耻 | 11 | 8342 | 803B | 耻 | | |
| 4090 | 脚 | 180 | 8347 | 811A | 脚 | | |
| 4104 | 莅 | 1 | 8355 | 8385 | 莅 | | |
| 4107 | 葱 | 7 | 8358 | 8471 | 葱 | | |
| 4111 | 蕤 | 1 | 835C | 8493 | 蕤 | | |
| 4129 | 虬 | 4 | 836E | 866C | 虬 | | |
| 4130 | 衅 | 4 | 836F | 8845 | 衅 | | |
| 4131 | 劔 | 1 | 8370 | 8842 | 劔 | | |
| 4134 | 衛 | 8 | 8373 | 885E | 衛 | | 衛，異體字問題。 |
| 4139 | 袴 | 1 | 8378 | 88B4 | 袴 | | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|--------------|---------------------|
| 4140 | 衽 | 1 | 8379 | 88B5 | 衽 | | |
| 4145 | 羈 | 2 | 837E | 898A | 羈 | | Unicode 字型 呈現差異。 |
| 4146 | 覩 | 1 | 83A1 | 89A9 | 覩 | | |
| 4167 | 賫 | 2 | 83B6 | 8CEB | 賫 | | |
| 4169 | 趁 | 1 | 83B8 | 8DA6 | 趁 | | |
| 4173 | 駝 | 7 | 83BC | 8EAD | 駝 | | |
| 4184 | 迹 | 44 | 83C7 | 8FF9 | 迹 | | |
| 4193 | 鈎 | 8 | 83D0 | 920E | 鈎 | | |
| 4195 | 鉢 | 4 | 83D2 | 9262 | 鉢 | | |
| 4218 | 菲 | 4 | 83E9 | 97EE | 菲 | | |
| 4247 | 鮎 | 1 | 8447 | 9B8E | 鮎 | | |
| 4256 | 鷄 | 60 | 8450 | 9DC4 | 鷄 | | |
| 4262 | 麪 | 6 | 8456 | 9EAA | 麪 | | Unicode 字型 呈現差異。 |
| 4267 | 鼈 | 1 | 845B | 9F08 | 鼈 | | |
| 4297 | 躔 | 1 | 8479 | 281E0 | | 𠃉 疋 厂 墨 𠃉 | |
| 4307 | 罵 | 48 | 84A5 | 99E1 | 罵 | | 罵，異體字問 |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|--------------|----------|
| | | | | | | | 題。 |
| 4308 | 刼 | 9 | 84A6 | 5227 | 刼 | | |
| 4316 | 艷 | 3 | 84AE | 8276 | 艷 | | |
| 4318 | | 1 | 84B0 | | | 形 畝且 毛 口 | 艷，異體字問題。 |
| 4345 | | 3 | 84CB | | | 口 𠂔 救 | 待造字。 |
| 4372 | | 15 | 84E6 | | | 形 广 回 禾 口 | 待造字。 |
| 4432 | | 55 | 8563 | | | 形 纟 厂 墨 口 | 纏，異體字問題。 |
| 4447 | 捏 | 33 | 8572 | 63D1 | 捏 | | |
| 4448 | | 1 | 8573 | | | 形 纟 厂 墨 口 | 纏，異體字問題。 |
| 4476 | | 1 | 85B1 | | | 良 𠂔 尸 | 即，異體字問題。 |
| 4518 | 响 | 1 | 85DB | 54CD | 响 | | |
| 4531 | 攢 | 1 | 85E8 | 6505 | 攢 | | |
| 4616 | 壳 | 1 | 867E | 58F3 | 壳 | | 壳，異體字問題。 |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|---------|----------|
| 4742 | | 2 | 875F | | | 衣△執 | 褻，異體字問題。 |
| 4761 | 標 | 2 | 8772 | 6AA9 | 標 | | |
| 4768 | 妬 | 1 | 8779 | 59AC | 妬 | | |
| 4796 | | 1 | 87B7 | | | 疒疒 豕 | 瘰，異體字問題。 |
| 4805 | 瀆 | 1 | 87C0 | 51DF | 瀆 | | 瀆，異體字問題。 |
| 4940 | 韵 | 17 | 88AA | 97F5 | 韵 | | |
| 5005 | 幫 | 3 | 88EB | 5E2E | 幫 | | |
| 5008 | 幪 | 3 | 88EE | 5E5E | 幪 | | |
| 5046 | 彝 | 2 | 8955 | 5F5C | 彝 | | |
| 5101 | 馱 | 9 | 89AE | 4B7E | | 馬△犬 | |
| 5163 | 憂 | 3 | 89EC | 621E | 憂 | | |
| 5218 | | 1 | 8A64 | | | 宀 、 | 寬，異體字問題。 |
| 5290 | 携 | 45 | 8ACE | 643A | 携 | | |
| 5299 | 戟 | 4 | 8AD7 | 39B8 | | 卓△戈 | |
| 5316 | 晋 | 1 | 8AE8 | 664B | 晋 | | 晉，異體字問題 |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-----|------|---------|-------|-----|----------|
| | | | | | | | 題。 |
| 5543 | 養 | 2 | 8C6F | 9B9D | 養 | | |
| 5551 | 費 | 1 | 8C77 | 8CF7 | 費 | | |
| 5590 | 朶 | 3 | 8CC0 | 6736 | 朶 | | |
| 5591 | 杞 | 3 | 8CC1 | 233CC | | 木𠂇巳 | 杞，異體字問題。 |
| 5621 | 泪 | 23 | 8CDF | 6CEA | 泪 | | |
| 5702 | 厨 | 123 | 8D71 | 53A8 | 厨 | | |
| 5705 | 廐 | 1 | 8D74 | 5ED0 | 廐 | | |
| 5706 | 敘 | 55 | 8D75 | 654D | 敘 | | 敘，異體字問題。 |
| 5742 | 澁 | 1 | 8DBB | 6F81 | 澁 | | |
| 5753 | 潜 | 279 | 8DC6 | 6F5C | 潜 | | |
| 5768 | 潤 | 4 | 8DD5 | 6FF6 | 潤 | | |

二、手動取代表說明：

《歧路燈》檔案中出現「●」字形，共計74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《歧路燈》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、《歧路燈》「●」字手動取代表

| 頁碼 | 檔案原文摘錄 | 原書字形 | 備註 |
|--------|---------|------|-------|
| p. 35 | 或閒●一韵 | 門△龟 | 改為門△龜 |
| p. 124 | 舌凍蟬● | 胃△力 | 改為勛 |
| p. 224 | 台公●孫 | 冢 | |
| p. 247 | 一個小● | 錦 | |
| p. 290 | 璀●奪目 | 灿 | 改為燦 |
| p. 296 | 朱仙鎮●裁刀 | 炯 | |
| p. 302 | 茅拔茹●箱 | 賴 | |
| p. 353 | ●孩子 | 尸△求 | 有 3 處 |
| p. 355 | 賒賬俺們● | 貝△青 | |
| p. 356 | 咱●賒賬哩 | 貝△青 | |
| p. 373 | 還沒福●哩 | 貝△青 | |
| p. 403 | 麵合●條子 | 餡 | |
| p. 407 | 只是一個● | 甬△毛 | |
| p. 411 | 廚房把鐵● | 鉞 | |
| p. 416 | ●症 | 瘳 | 有 2 處 |
| p. 438 | ●去。」譚紹聞 | 貝△青 | |
| p. 439 | 只●我的頭 | 貝△青 | |

| 頁碼 | 檔案原文摘錄 | 原書字形 | 備註 |
|--------|--------|------|-------|
| p. 441 | 小字●兒 | 滙 | |
| p. 442 | 一總●明 | 算 | |
| p. 580 | 是烏龜●了 | 貝厶青 | |
| p. 587 | 身披著●毛 | 角厶毛 | |
| p. 604 | 搬了一個●子 | 橈 | |
| p. 604 | 蹬翻小馬● | 橈 | |
| p. 604 | 扶起●子 | 橈 | |
| p. 669 | 狗● | 𠂔 | 有 7 處 |
| p. 670 | ●兒，媳婦兒 | 𠂔 | |
| p. 681 | 奇殃，●創 | 懲 | |
| p. 683 | 騰●齊聲 | 雲 | |
| p. 726 | 或●勞金 | 貝厶青 | |
| p. 758 | ●庠 | 胶 | 有 2 處 |
| p. 765 | 踊● | 跃 | 有 2 處 |
| p. 770 | 不得踏确● | 𠂔厶牢 | 待造字 |
| p. 793 | 放個●子坐下 | 橈 | |
| p. 793 | 放在●子上 | 橈 | |
| p. 798 | ●耀奪目 | 灿 | 改為燦 |
| p. 804 | 如●不中用 | 狷 | |
| p. 809 | 與相公●了 | 淒 | |
| p. 809 | 下樓來吃●柿 | 淒 | |

| 頁碼 | 檔案原文摘錄 | 原書字形 | 備註 |
|---------|--------|------|-----|
| p. 829 | 是瞎●話 | 尸△求 | |
| p. 831 | 瓊漿● | 卣 | |
| p. 835 | ●理陰陽 | 燮 | |
| p. 838 | 喇叭鏡●齊響 | 金△友 | 待造字 |
| p. 853 | 集債如● | 狷 | |
| p. 856 | 都●而拔之 | 艹△娠 | |
| p. 938 | 早已形縮如● | 狷 | |
| p. 944 | 這養娘●婦門 | 爨 | |
| p. 975 | 各持絲● | 綵 | 改為縹 |
| p. 985 | 得侍●比 | 臬 | |
| p. 1007 | 圪● | 土△達 | 有3處 |
| p. 1060 | 必定●喝你 | 口△么 | |
| p. 1068 | 成千斤白● | 砢 | |
| p. 1068 | 紙上加● | 砢 | |
| p. 1068 | 一分白● | 砢 | |
| p. 1070 | 已矢集如● | 狷 | |
| p. 1075 | ●從誨以禮義 | 月△脅 | 待造字 |
| p. 1079 | 由他●喝官府 | 口△么 | |
| p. 1080 | 撫藩●的旗 | 自△本 | |
| p. 1086 | 扇袋，梳● | 篋 | 改為篋 |
| p. 1105 | 機●一架 | 箴 | |

| 頁碼 | 檔案原文摘錄 | 原書字形 | 備註 |
|---------|--------|------|-----|
| p. 1107 | 不敏●陋 | 謏 | |
| p. 1108 | 吃了合●盞 | 承△乙 | 改為盞 |