

《商君書》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
2008/2/15 黃榮順 製作

一、《商君書》一書(商君書.xml)使用舊版造字 8 個，字頻 10 次，詳如附件一。這 8 個造字中，8 個可轉成 Windows XP 可顯示的字，字頻 10 次；另外 0 個字必須轉成構字式，字頻 0 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 4115 的「藁」字，原字應為「藁」。

三、「商君書.xml」發現之未確定字形「●」共 4 處，修正情形詳列附件二，附件二「商君書.xml」「●」修改表體例說明如下：

甲、檔案原文摘錄：摘錄檔案(商君書.xml)中包含「●」的文句段落，其中紅字底線的部分，為需要手動取代的原文。

乙、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、「商君書.xml」造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3804	効	1	1	8163	52B9	効		
3852	嫵	1	1	81B5	5A63	嫵		
3883	徧	1	1	81D4	5FA7	徧		
3888	惠	3	3	81D9	60B3	惠		
4115	藁	1	1	8360	85C1	藁		藁，舊漢籍錯字，可用 程式全部取代。
4167	賁	1	1	83B6	8CEB	賁		
4242	鬪	1	1	8442	9B2D	鬪		
4768	妬	1	1	8779	59AC	妬		

附件二、「商君書.xml」「●」修改表

檔案原文摘錄	手動取代結果
令有甬，官食●	令有甬，官食既△本
宛鉅鐵●，利若蜂	宛鉅鐵金△色，利若蜂
清朝日●，則上別飛鳥	清朝日黃△崙，則上別飛鳥
堯舜禹湯且皆如●焉	堯舜禹湯且皆如馬△致焉