

《老子》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
2008/1/28 黃榮順 製作

一、《老子》一書(老子.xml)使用舊版造字 22 個，字頻 84 次，詳如附件一。這 22 個造字中，17 個可轉成 Windows XP 可顯示的字，字頻 69 次；另外 5 個字必須轉成構字式，字頻 15 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、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 4134 的「衛」係「衛」字之異體，故以「衛」字取代。
- 2、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 5034 的「弃」字，Unicode 字型呈現為「弃」，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 5200 的「ノ門米大」。

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。所以如造字編號 5459 的「缺」字，Unicode 編碼為 21647，由於 Windows XP 並不支援，仍須改用構字式「垂夬」。

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、「老子.xml」發現之未確定字形「●」共 37 處，修正情形詳列附件二，附件二「老子.xml」未確定字形「●」修改表體例說明如下：

- 甲、檔案原文摘錄：摘錄檔案(老子.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。
- 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。古漢字部份尚待處理，則維持紅字底線的形式。

附件一、「老子.xml」造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1396	罕	1	1	91ED	7A82	罕		
2044	踈	1	1	9642	8E08	踈		
3789	冲	2	2	8154	51B2	冲		
3814	却	1	1	816D	5374	却		
3944	氷	1	1	8252	6C37	氷		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3996	疎	1	1	82A8	758E	疎		
4104	莅	1	1	8355	8385	莅		
4132		9	9	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	1	1	8373	885E	衛		衛，異體字問題。
4160	適	1	1	83AF	8B81	適		
4172	屨	1	1	83BB	8E77	屨		
4184	迹	1	1	83C7	8FF9	迹		
4203	隣	1	1	83DA	96A3	隣		
4256	鷄	1	1	8450	9DC4	鷄		
4455	劒	1	1	857A			僉𠂇刀	
4519	咲	2	2	85DC	54B2	咲		
4907	橐	1	1	8867	69D6	橐		
5034	弃	1	1	8949	5F03	弃		Unicode 字型差異。
5200		1	1	8A52			𠂇ノ門米大𠂇	待造字。
5459	缺	3	3	8BDA	21647		垂𠂇夂	
5749		1	1	8DC2			片𠂇庸	待造字。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5860	无	51	51	C6D3	65E0	无		

附件二、「老子.xml」未確定字形「●」修改表

檔案原文摘錄	手動取代結果
以萬物爲●●	以萬物爲𠃉++マ三☐牙△苟
以百姓爲●●	以百姓爲𠃉++マ三☐牙△苟
富貴而驕，自遺其●	富貴而驕，自遺其各△、
五音令人耳●	五音令人耳𠃉立月巳耳☐
致虛極，守靜●	致虛極，守靜++△馬
●冥中有精	穴△幻冥中有精
●者不行	大△于者不行
或強或羸，或接或●	或強或羸，或接或隳
若美之，是樂●人	若美之，是樂𠃉クマ女心☐人
夫樂●者	夫樂𠃉クマ女心☐者
●人眾多	𠃉クマ女心☐人眾多
多●必厚亡	多茂△臣必厚亡
却走馬以●	却走馬以米△田△土
●莫大於欲得	各△、莫大於欲得
陸行不遇●●	陸行不遇坐△儿虐
●無所投其角	坐△儿無所投其角
●無所措其爪	虐無所措其爪
是謂盜●	是謂盜大△于
毒●不螫	毒ノ△虫不螫
骨弱●柔而握固	骨弱++△勍柔而握固
雖有拱●以先駟馬	雖有拱辟△王以先駟馬
●難於易	置難於易
治之於未●	治之於未舌△L
●者失之	𠃉卉一未丸☐者失之

檔案原文摘錄	手動取代結果
無●，故無失	無 [○] 土 [○] 未 [○] 丸 [○] ，故無失
仍无敵，●无兵	仍无敵， [○] 弁 [○] 一 [○] 未 [○] 丸 [○] 无兵
輕敵幾●吾寶	輕敵幾 [○] 亡 [○] 吾寶
吾●得而●之	吾 [○] 土 [○] 未 [○] 丸 [○] 得而 [○] 心 [○] 之
常有司●者●	常有司 [○] 者 [○]
夫代司●者	夫代司 [○] 者
是謂代大匠●	是謂代大匠 [○]
夫代大匠●	夫代大匠 [○]