

《醫方類聚》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
2008/7/4 黃榮順 製作

一、《醫方類聚》一書(醫方類聚.xml)使用舊版造字 35 個，字頻 2258 次，詳如附件一。這 35 個造字中，15 個可轉成 Windows XP 可顯示的字，字頻 2188 次；另外 20 個字必須轉成構字式，字頻 70 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，必須人工刪除：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，雖然頻次相同，其為原書所無之誤增字，仍須手動刪除。如編號 3360 的「」（造字區未定義），原書所無，為誤增字，手動刪除。
- 2、校對問題，新舊版缺字並存，原書所無，配合刪除：漢籍在舊轉新的過程中，以人工修改方式取代舊漢籍造字，雖然修改的結果是正確的，但是只修改到部分造字，沒有全部修改到，導致舊漢籍造字在新版漢籍檔案中，一部分已被取代成 Unicode 或構字式，另一部分仍為舊漢籍造字的並存情形。處理方式為配合漢籍工作室的修改，將沒被修改到的舊漢籍造字，以對應的 Unicode 或構字式取代。如編號 3716 為「」，舊漢籍造字區未定義此字碼，比對 txt 檔及 xml 檔後，發現原先在 txt 檔使用此字碼的位置皆為誤增字，原書所無，xml 檔中部份頻次已刪除，故配合修改，將編號 3716 其餘未修改的頻次手動刪除。

- 3、重碼問題，採用部件外字，不轉碼：舊漢籍造字與構字式部件外字，皆是建立在 Big5 的使用者造字區中，兩者所使用的碼區是重疊的，在同時使用構字式和舊漢籍造字的新版漢籍中，這些重疊的字碼可能被用以表達兩個不同字形，我們稱之為「重碼問題」。然而漢籍在舊轉新的過程中，若已將原舊漢籍造字的部分全數改為 Unicode 或構字式，則在 xml 出現的造字頻次皆採用部件外字，此時便不需轉碼。如編號 5338，字碼為 8AFE，對應的舊漢籍造字為「暎」，部件外字為「艹」，對原書發現書中皆取部件外字「艹」，故不需轉碼。
- 4、重碼問題，採用部件外字，未確定字形，手動取代：舊漢籍造字與構字式部件外字，皆是建立在 Big5 的使用者造字區中，兩者所使用的碼區是重疊的，在同時使用構字式和舊漢籍造字的新版漢籍中，這些重疊的字碼可能被用以表達兩個不同字形，我們稱之為「重碼問題」。在舊轉新的過程中，若已將原舊漢籍造字的部分全數改為 Unicode 或構字式，則在 xml 出現的造字頻次皆採用部件外字，此時便不需轉碼。但部件外字為未定義字形時，則須以原書字形手動取代。如編號 5809，舊漢籍造字字形為「踰」，部件外字為「⊙」（未定義字形），必須一一比對原書，加以手動取代，取代表參考附件三。

三、 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：舊漢籍造字。
- 丙、 取代：使用字形與造字欄不同時，於此欄列出。
- 丁、 檔案原文摘錄：摘錄檔案(醫方類聚.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。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。
- 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。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。

四、「醫方類聚.xml」發現之未確定字形「」共 11 處，修正情形詳列附件三，附件三「醫方類聚.xml」修改表體例說明如下：

甲、檔案原文摘錄：摘錄檔案(醫方類聚.xml)中包含「」未確定字形的文句段落，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。

五、「醫方類聚.xml」發現之缺字外錯字共 8 處，修正情形詳列附件五，附件五「醫方類聚.xml」缺字外錯字修改表體例說明如下：

甲、檔案原文摘錄：摘錄檔案(醫方類聚.xml)中包含該錯字的文句段落，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。

丙、備註：處理情形需要特別說明時，則於此欄註出。

附件一、「醫方類聚.xml」造字分析表

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-----|--------------|
| 3360 | | 6 | 6 | 9E7E | | | | 舊版漢籍錯字，手動刪除。 |
| 3362 | | 1 | 1 | 9EA2 | | | | 舊版漢籍錯字，手動刪除。 |
| 3389 | | 1 | 1 | 9EBD | | | | 舊版漢籍錯字，手動刪除。 |
| 3442 | | 1 | 1 | 9EF2 | | | | 舊版漢籍錯字，手動刪除。 |
| 3456 | | 1 | 1 | 9F41 | | | | 舊版漢籍錯字，手動刪除。 |

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-------|--------------------|
| 3459 | | 1 | 1 | 9F44 | | | | 舊版漢籍錯字，手動刪除。 |
| 3480 | | 1 | 1 | 9F59 | | | | 舊版漢籍錯字，手動刪除。 |
| 3497 | | 1 | 1 | 9F6A | | | | 舊版漢籍錯字，手動刪除。 |
| 3516 | | 1 | 1 | 9F7D | | | | 舊版漢籍錯字，手動刪除。 |
| 3528 | | 1 | 1 | 9FAB | | | | 𠄎，重碼問題，採用部件外字，不轉碼。 |
| 3618 | | 1 | 1 | A046 | | | | 舊版漢籍錯字，手動刪除。 |
| 3698 | | 1 | 1 | A0B8 | | | | 舊版漢籍錯字，手動刪除。 |
| 3716 | | 16 | 13 | A0CA | | | | 舊版漢籍錯字，手動刪除。 |
| 3757 | | 3 | 3 | A0F3 | | | | 舊版漢籍錯字，手動刪除。 |
| 4278 | 亘 | 12 | 2 | 8466 | 4E98 | 亘 | | 𠄎，重碼問題，採用部件外字，不轉碼。 |
| 4904 | 尅 | 2 | 2 | 8864 | 5C05 | 尅 | | 𠄎，重碼問題，採用部件外字，不轉碼。 |
| 5009 | 廠 | 1 | 3 | 88EF | 222C5 | | 广 𠄎 歛 | 舌，重碼問題，採用部件外字，不轉碼。 |
| 5335 | 𧈧 | 1 | 28 | 8AFB | 459F | | 亡 𧈧 虫 | 𧈧，重碼問題，採用部件外字，不轉碼。 |
| 5338 | 曠 | 1 | 548 | 8AFE | 668E | 曠 | | 𠄎，重碼問題，採用部件外 |

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-----|--------------------|
| | | | | | | | | 字，不轉碼。 |
| 5344 | 磬 | 2 | 1 | 8B45 | 55F8 | 磬 | | 久，重碼問題，採用部件外字，不轉碼。 |
| 5347 | 憾 | 1 | 5 | 8B48 | 617D | 憾 | | 𠂇，重碼問題，採用部件外字，不轉碼。 |
| 5353 | 役 | 2 | 303 | 8B4E | 4F07 | 役 | | 主，重碼問題，採用部件外字，不轉碼。 |
| 5386 | 埆 | 1 | 1 | 8B6F | 365B | | 土𠂇專 | 开，重碼問題，採用部件外字，不轉碼。 |
| 5408 | 紕 | 9 | 1 | 8BA7 | 7D4D | 紕 | | 亲，重碼問題，採用部件外字，不轉碼。 |
| 5445 | 覘 | 1 | 1269 | 8BCC | 6898 | 覘 | | 疒，重碼問題，採用部件外字，不轉碼。 |
| 5448 | 髑 | 11 | 40 | 8BCF | 9AC3 | 髑 | | 音，重碼問題，採用部件外字，不轉碼。 |
| 5482 | 璽 | 7 | 1 | 8BF1 | 8812 | 璽 | | 丰，重碼問題，採用部件外字，不轉碼。 |
| 5493 | 之 | 1 | 1 | 8BFC | | | 之 | 彡，重碼問題，採用部件外字，不轉碼。 |
| 5558 | 龕 | 1 | 2 | 8C7E | 9E84 | 龕 | | 𠂇，重碼問題，採用部件外字，不轉碼。 |
| 5578 | 弛 | 2 | 3 | 8CB4 | 38AE | | 弓𠂇屯 | 辰，重碼問題，採用部件外字，不轉碼。 |

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-----|--------------------------------------|
| 5583 | 殮 | 2 | 1 | 8CB9 | 98F1 | 殮 | | 殮，重碼問題，採用部件外字，不轉碼。 |
| 5588 | 恡 | 4 | 1 | 8CBE | 6061 | 恡 | | 恡，重碼問題，採用部件外字，不轉碼。 |
| 5769 | 邠 | 2 | 1 | 8DD6 | 90CD | 邠 | | 邠，重碼問題，採用部件外字，不轉碼。 |
| 5809 | 踰 | 13 | 11 | 8DFE | 8E79 | 踰 | | 重碼問題，採用部件外字，使用未定義字形②，故以原書字形取代，參考附件三。 |
| 5815 | | 1 | 1 | C6A6 | | | | 舊版漢籍錯字，手動刪除。 |

附件二、「醫方類聚.xml」手動取代表

| 編號 | 造字 | 檔案原文摘錄 | 手動取代結果 |
|------|----|---------------------|------------|
| 3360 | | <u>𪗇</u> （手動全部刪除） | （手動全部刪除） |
| 3362 | | <u>𪗇</u> □□及脈平無証者 | □□及脈平無証者 |
| 3389 | | 學士曰：金匱 <u>𪗇</u> | 學士曰：金匱 |
| 3442 | | <u>𪗇</u> 牡蠣<qe>[二]蠣 | 牡蠣<qe>[二]蠣 |
| 3456 | | ，癢癢不時， <u>𪗇</u> | ，癢癢不時， |
| 3459 | | </f> 礪砂 廣求 <u>𪗇</u> | </f> 礪砂 廣求 |
| 3480 | | 苦參 豬牙皂 <u>𪗇</u> | 苦參 豬牙皂 |

| 編號 | 造字 | 檔案原文摘錄 | 手動取代結果 |
|------|----|-----------------------|--------------|
| 3497 | | <u>点</u> 胡黃連<f s=" | 胡黃連<f s=" |
| 3516 | | 穀大，每服五、七 <u>賜</u> | 穀大，每服五、七 |
| 3618 | | 瘡裏有蟲癢痛，搔之 <u>俞</u> | 瘡裏有蟲癢痛，搔之 |
| 3698 | | 、粳米大</f> <u>巴</u> | 、粳米大</f> 巴 |
| 3716 | | <u>和</u> （手動全部刪除） | （手動全部刪除） |
| 3757 | | 黃耆 甘草 梔 <u>箕</u> | 黃耆 甘草 梔 |
| | | 防風 <u>箕</u> 黃耆 | 防風 黃耆 |
| | | <u>箕</u> 巴豆<f s="-2"> | 巴豆<f s="-2"> |
| 5815 | | 卷一百二臟腑門__門 | 卷一百二臟腑門 |

附件三、「醫方類聚.xml」 「❓」修改表

| 檔案原文摘錄 | 手動取代結果 |
|-----------|--------------------|
| ❓化痰涎 | <u>善</u> 化痰涎 |
| ❓醫療，與藥服之 | <u>善</u> 醫療，與藥服之 |
| ❓，帶好生犀角佳也 | <u>善</u> ，帶好生犀角佳也 |
| 餘者猶❓爾 | 餘者猶 <u>善</u> 爾 |
| 即❓，治療蛇毒亦效 | 即 <u>善</u> ，治療蛇毒亦效 |
| 皆微候，宜❓察之 | 皆微候，宜 <u>善</u> 察之 |
| 兩乾而成坤❓命也 | 兩乾而成坤 <u>☷</u> 命也 |
| ❓離中之陰，乃真精 | <u>☷</u> 離中之陰，乃真精 |
| 心屬離❓ | 心屬離 <u>☲</u> |
| 後獨❓，解音樂誦 | 後獨 <u>善</u> ，解音樂誦 |
| 只好白石☶❓土亦 | 只好白石☶ <u>善</u> 土亦 |

附件四、「醫方類聚.xml」缺字外錯字取代表

| 檔案原文摘錄 | 手動取代結果 | 備註 |
|----------------------|----------------------|----|
| 楊一者， <u>??</u> | 楊一者， | 刪除 |
| 鴨亦 <u>??</u> | 鴨亦 | 刪除 |
| ，肢節煩痛方 <u>?</u> </s> | ，肢節煩痛方 <u>:</u> </s> | |
| 卷一百二臟腑門 <u>門</u> | 卷一百二臟腑門 | 刪除 |
| 父為蚯蚓 <u>Δ??</u> 畱蟲 | 父為蚯蚓 <u>Δ</u> 畱蟲 | 刪除 |