

中文電腦缺字解決方案

漢字構形資料庫
使用手冊

中央研究院 資訊科學研究所 文獻處理實驗室
教育部 圖書館自動化及網路化策略推動委員會

民國九十一年七月

漢字構形資料庫使用手冊 2002 年 7 月

漢字庫字型為葉健欣開發

北師大小篆字型為北京師範大學研發，中央研究院 資訊科學研究所 文獻處理實驗室重新編碼。

Java Applet 開發的程式為中央研究院 資訊科學研究所 網路與通訊實驗室開發

《漢語大字典》為遠東圖書公司發行

委辦單位：教育部 圖書館自動化及網路化策略推動委員會

執行單位：中央研究院 資訊科學研究所 文獻處理實驗室

計畫主持人：謝清俊

編著：莊德明 許婉蓉

編輯：張翠玲

審訂：謝瀛春

出版：中央研究院 資訊科學研究所 文獻處理實驗室

2000 年 12 月 第一版

2002 年 7 月 第二版

ISBN：957-671-871-6

非賣品

115 台北市南港區研究院路二段 128 號

中央研究院 資訊科學研究所

網址：<http://www.sinica.edu.tw/~cdp>

目 次

序	i
內容簡介	v
第一章 系統安裝及移除	1
一、系統需求	1
二、安裝系統	2
三、移除系統	7
第二章 系統功能簡介	8
一、使用 Microsoft Office 處理缺字	9
缺字轉換次功能表	9
漢字構形資料庫使用界面功能表列	9
構字符號工具列	10
在 Word 處理缺字	12
編輯文件內容	20
構字式和一般文字	22
構字輸入法	22
在 Excel 處理缺字	22
在 PowerPoint 處理缺字	24
在 FrontPage 處理缺字	25

二、漢字構形資料庫使用界面	28
開啟漢字構形資料庫使用界面	28
查詢字形（輸入單一部件）	29
查詢字形（輸入多個部件）	31
複製構字式	32
查詢異體字	33
查詢字形源流演變	34
查詢字形結構	35
三、網際網路的缺字處理	38
網頁的缺字表達	38
使用 Java Applet 轉換網頁中的構字式	40
如何在網頁加入 Java Applet	41
使用 Microsoft Office 處理網頁缺字	44
在網際網路上查詢缺字與異體字	45
第三章 構字式的自動校正	47
一、一般文字的構字式	47
二、同一字形的不同構字式	47
三、字形相似的構字式	50
四、包含異寫部件的構字式	51
第四章 缺字字型	53

一、支援單一字型	53
二、複製缺字構字式	54
三、古文字字型	55
第五章 實例輸入	57
一、文件輸入須符合的條件	57
二、缺字的輸入	59
安裝漢字構形資料庫	59
使用 Word 來輸入《說文》釋「教」的段落	59
三、字形轉換	75
四、構字式自動校正	76
附錄一 缺字的表達方式	79
一、部件	79
二、構字符號	80
三、構字式	81
1. 橫連	81
2. 直連	81
3. 包含	81
4. 部件序	82
5. 方便符號	83

序

在 1994 的時候，中央研究院數位化的古籍已有相當的數量，因之而累積的缺字也已超過了八千個，形成了相當嚴重的問題。缺字是由於漢字字碼中的字形不足引起的。資料庫有了缺字，不僅影響到使用者對古籍的閱覽，更影響到對古籍的了解、檢索、列印，以及後續的種種處理，包括資料的交換和共享等問題。有鑑於此，本實驗室開始設法解決此一問題。

解決缺字問題的根本，在於解決現行漢字交換碼的根本缺失。漢字的字集是一個開放性質的；也就是說，漢字的字集依古今的變異、專業與應用環境的差異等而有字數、字形、字音以及字義上的變化。音義且不談，僅就字數而言，即已不適合作固定數量的限定；這與數量已定的西方語言的『字母集』，是不可以一概而論的。然而，現行漢字交換碼的結構，卻仿照西方語言的字母集的結構來設計，這不能不說是『削足適履』。

再者，語言是有生命的，任何活在當下的語言，都需要有創作新詞（word）的能力，否則必定無法配合社會變遷之需，而終將面臨被淘汰的命運。我們不時看見有英文新詞（字），就是個很好的例子。可是，在電腦中限定漢字的字數，就斬斷了漢語言電腦中創作新字（詞）的生命力，埋下了日後不能適用的因子。這個問題也不應等閒視之。

我們解決缺字問題的方法，是遵從漢字構形的原理，

附錄二 構字符號的輸入	84
附錄三 常用部件外字的輸入	85
附錄四 安裝部件外字及輸入法	88
重新安裝部件外字及輸入法	88
恢復使用者原先的外字	91

對漢字字形的結構做制式表達與詳細的分析。一個字的字形結構式，是該字極佳的識別符號與工具；因為字形若不一樣，則字形結構必不相同；反之，字形結構若相同，其形亦必同。事實上，漢字字形結構的拆分與分析，淵源流長。撇開文字學的研究不談，最早為電腦拆分漢字的研究，見於 1972 年交通大學關於『交大字根系統』一系列的論文和報告。由於交大字根系統的相關電子檔案已不能讀出，於是在本研究剛開始的時候，我們就決定沿用交大字根系統的方法重新再做一遍，並在過程中作若干的改進。這段工作最大的工程，就是建立了『漢字構形資料庫』。

爲了建構漢字構形資料庫，我們徹底檢查了《漢語大字典》、《說文解字》和一些交換碼與各國所頒布的字集標準，分析了其中的每一個字。所以，目前建構的系統並不僅僅是做構字的表達與識別，還做了兩件大事：一是古今字形與古今構字的銜接和對映，二是異體字形的表達和處理。這都是電腦裡多年來懸而未決的問題。

在古今文字銜接和對映的部份，我們在中央研究院歷史語言研究所古文字組的指導和協助下，已逐字分析了《說文解字》中所有的字形，建構了小篆字形的構形資料庫，並且已和『漢字構形資料庫』做了對映與聯繫。所以，從我們的系統，使用者可以從現代字形直接聯繫到小篆與楷化的小篆字形，也可由小篆的構形直接查核現代字形。換言之，使用者可不必具有任何文字學的知識，即可查閱到相當多古文字相關的知識和資料。我們相信，這對漢語的教學和應用是有幫助的。至於小篆以前的字形，主要是金文（鐘鼎文）和甲骨文。金文的構形資料庫部份將於半年內完成，並將與歷史語言研究所古文字組所做的『金文資

料庫』銜接和對映。甲骨文部份則尚待努力。

在異體字形方面，我們的系統中已將《漢語大字典》中一萬兩千組以上的異體字表建構成了異體字資料庫。此外，系統也建構了異體字根與異體部件的對映關係，以便使用者能更人性化的使用本系統。我們相信，異體字資料庫能協助我們解決久懸未決的異體字問題，特別是在資料檢索方面，可以提高檢索的精確度和召回率。

由於本系統的設計理念與目前的交換碼完全不同，所以本系統可以與使用任何交換碼的系統相容。換言之，任何亞東文字的處理系統，無論是中日韓越、無論是簡繁，都可以附加本系統作為徹底解決缺字的機制。目前我們推出的是基於五大碼（Big-5）的系統，可將五大碼能處理的約一萬三千字形，立刻擴充到六萬以上，並擴及小篆和異體字等。更重要的是，這能力的擴大，並不需要佔用原交換碼的編碼空間，也不需要修改原來的應用軟體！

本系統的發展，歷時七年餘。此間得中央研究院資訊科學研究所的信賴與支持最多。如果資訊科學研究所對本實驗室的信賴不足，是不可能讓我們花七年多的時間，來做個這麼繁重冗長的工作。在此特別致謝。此外，爲了要開發在網路與 WINDOW 下的應用系統，教育部（圖書館自動化及網路化策略推動委員會）與國科會（數位博物館和數位典藏計畫）在最近兩年的支援，也有雪中送炭、臨門一腳的效果。這也是我們要特別感謝的。在系統工程方面，主要的工作人員是莊德明先生和許永成先生。他們兩位任勞任怨和不眠不休的毅力，是使得本計畫能堅持到最後一刻的關鍵，也是推動本計畫的主要動力。曾在本實驗

室工作的所有人員也是應感謝的，因為在我們忙得不可開交的時候，總是他們挽起袖子來幫忙。此外，在此七年中，本計畫與新竹科學園區的易符公司經常相互切磋、交換心得，是一段令人懷念的愉快時光，值得一記。

最後，本實驗室謹以謙誠的心，將此系統免費地獻給所有的同胞，願您使用本系統獲得利益！畢竟，文字是我們文化的根，能在這個根上盡一份力量，已是我們最大的福報。

謝清俊

中華民國九十一年十月

於資訊科學研究所文獻處理實驗室

內容簡介

使用電腦來處理漢字資料時，常會遇到一些電腦沒有的字形，以致無法輸入，這便是缺字問題。尤其是在處理古代文獻時，缺字的情況特別嚴重。面對缺字問題，「漢字構形資料庫」可說是一個成本較低、功能又較完備的解決方案。

漢字構形資料庫是利用漢字的構形來表達缺字，這個表達方式稱為「構字式」。除了構字式外，資料庫中還登錄了漢字的相關屬性，如筆畫、部首、字音、字型等。

這本手冊在介紹如何在 Microsoft Office 或網際網路中利用漢字構形資料庫來處理缺字，包含缺字的輸入、查詢、顯示、列印及交換。

手冊共分五章及四個附錄。第一章敘述系統的安裝及移除。第二章舉例說明系統的基本操作方式，包括構字式的輸入、缺字的列印及查詢。第三、四兩章指出系統的一些特別的功能和限制。第五章為實例解說。

附錄一說明缺字的表達方式，附錄二、三、四分別說明構字符號的輸入、常用部件外字的輸入以及安裝部件外字及輸入法。

第一章 系統安裝及移除

本章說明系統適用的作業環境、需求及限制。協助使用者安裝及移除系統，並對安裝時可能出現的訊息或安裝錯誤的原因及資訊，提供明確的解說。

一、系統需求

欲使用漢字構形資料庫及 Microsoft Office 增益集，必須有以下的配備

- 個人電腦：IBM 或 100%相容的個人電腦
- 作業系統：中文視窗 95/98/Me/2000/XP
- CPU：486 以上
- 光碟機：兩倍速以上
- 主記憶體：32MB 以上
- 硬碟：50MB 空間
- 應用軟體：Microsoft Office 2000/XP 中文版

二、安裝系統

1. 打開電腦的電源開關，啟動中文視窗作業系統。
2. 將「漢字構形資料庫」光碟片放入光碟機。

螢幕上會出現：

確定要安裝漢字構形資料庫？

3. 若要安裝，請按「確定」，否則按「取消」。

若未出現上述的訊息，可選擇下面 1 或 2 任一種方式來安裝系統：

1. 按一下螢幕左下方的「開始」，選擇「執行(R)」，在指令列中，請鍵入 D:\setup，按「Enter」(假設光碟機在 D 槽，若在 E 槽則鍵入 E:\setup，餘依此類推)。
2. 開啓「檔案總管」，在光碟機的根目錄內找到 setup.exe，按兩下開始執行安裝程式。

進入安裝程序後，畫面上會出現對話方塊：

安裝程式無法安裝使用中的系統檔案，也無法更新使用中的的共用檔案。建議您在安裝前，關閉所有的應用程式。

關閉其他的應用程式，然後按「確定」。若不想安裝，則按「結束安裝」。

按確定之後，出現以下畫面：

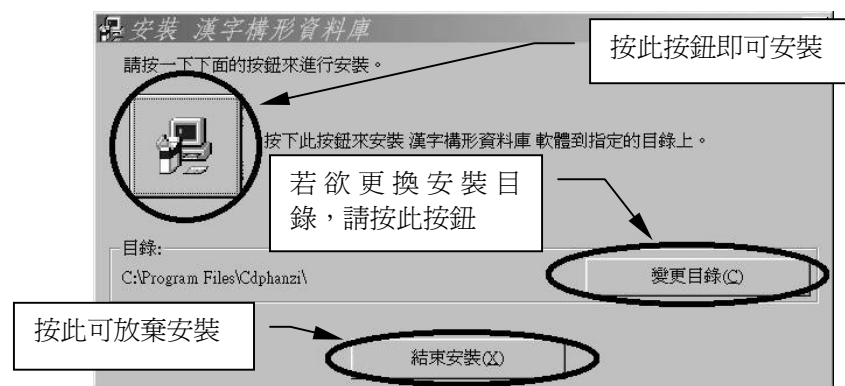


圖 1.1 安裝漢字構形資料庫

按一下最左上角的按鈕，開始安裝。

接下來，出現對話方塊：

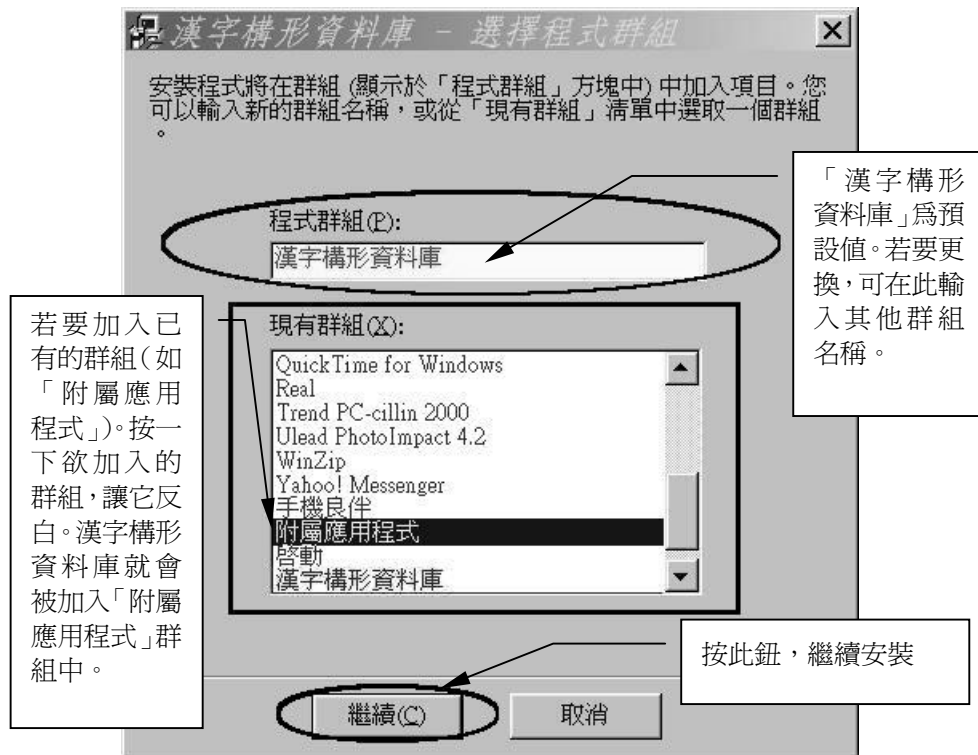


圖 1.2 選擇程式群組

按「繼續」後，安裝程式開始抄錄檔案，抄錄完成時，畫面出現下列訊息（見第 5 頁，圖 1.3）：

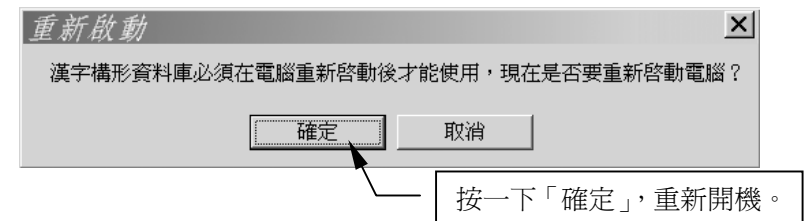


圖 1.3 重新啟動電腦對話框

按下「確定」，重新啟動電腦，即可完成安裝程序。

注意：

安裝過程中如果顯示以下訊息：

因為某些系統檔並非最新版，導致無法繼續安裝。假如您希望程式自動更新這些檔案，請選「確定」；您必須重新開機後，才能再執行安裝程式。假如您不希望變動這些檔案，請選擇「取消」。

選擇「確定」之後，繼續出現以下訊息：

要現在啟動視窗系統嗎？如果您選擇「否」，則必須等到下一次啟動後，才能再執行安裝程式。

此時請選擇「確定」，讓電腦重新開機。然後再重新執行一次光碟片上的 setup.exe。

選擇「取消」則離開安裝程序。

三、移除系統

依照下列程序移除系統

1. 按「開始」、「設定」、「控制台」
2. 按兩下「新增/移除程式」
3. 選擇「漢字構形資料庫」，按「新增/移除」
4. 按「確定」按鈕執行移除動作
5. 移除程式會再詢問您是否確定要移除，請按「確定」
6. 移除完畢後請按「確定」按鈕

注意：

1. 反安裝程式僅能刪除最初在安裝漢字構形資料庫時的一些相關檔案，因此，在反安裝過程中，若出現不能移除漢字構型資料庫目錄的訊息時，可能是使用者將檔案儲存在漢字構形資料庫的目錄中，導致目錄無法刪除。
使用者可以把這些檔案刪除或搬移到其他目錄，再刪除漢字構形資料庫的目錄。
2. 重新開機後再移除。

第二章 系統功能簡介

本章說明系統的基本功能及操作方式。包括使用 Microsoft Office 處理缺字、漢字構形資料庫使用界面的介紹及網際網路的缺字處理。系統架構圖如下圖所示：

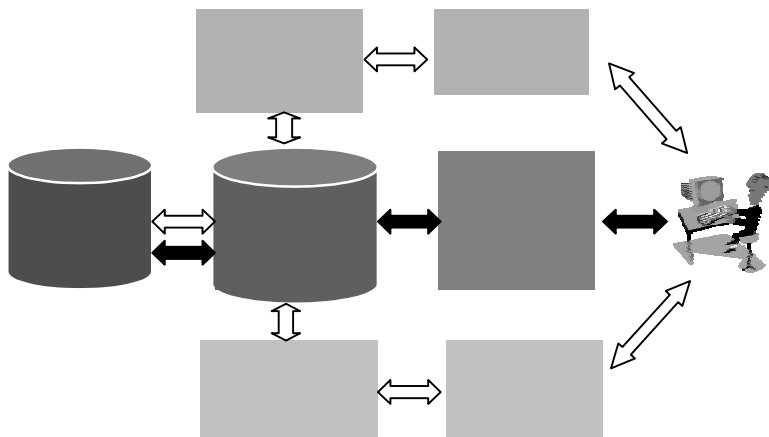


圖 2.1 漢字構形資料庫系統架構圖

圖 2.1 說明使用漢字構形資料庫有三種方式：

1. 利用本系統提供的 Microsoft Office 缺字處理增益集（詳見本章第一節）。
2. 直接使用漢字構形資料庫使用界面（詳見本章第二節）。
3. 使用網頁內含的 Java Applet（詳見本章第四節）。

缺字字型目前採用漢字庫字型，共有 12 套字型，包含 hzk1 至 hzkc。

一、使用 Microsoft Office 處理缺字

安裝系統會新增缺字處理增益集「cdpmsoai.dll」（dll 中的 l 為英文字母 L 的小寫），此增益集包含「缺字轉換次功能表」、「漢字構形資料庫使用界面」功能表列和「構字符號」工具列等使用界面

缺字轉換次功能表

按「工具」功能表、選擇「缺字轉換次功能表」中的各項功能，將缺字轉換成字形或圖片，或將圖片轉換成構字式。

漢字構形資料庫使用界面功能表列

按「工具」功能表、再按「漢字構形資料庫使用界面」，開啓漢字構形資料庫協助輸入缺字。

構字符號工具列

構字符號工具列可以協助使用者輸入缺字，在構字符號按鈕上按一下，構字符號就會加入文件當中。



圖 2.2 構字符號工具列

• 顯示構字符號工具列

若要顯示「構字符號」工具列，可選擇以下任一方式開啓構字符號工具列：

1. 按一下「檢視」功能表，選「工具列」，會跳出所有工具列的清單。將滑鼠移到最下面未打「✓」的「構字符號」上，按一下，即可顯示。(見第 10 頁，圖 2.3)
2. 將滑鼠移到任一工具列上，按一下滑鼠右鍵，會跳出所有工具列的清單。把滑鼠移到最下面未打「✓」的「構字符號」上，按一下，即可顯示。(見第 10 頁，圖 2.3)

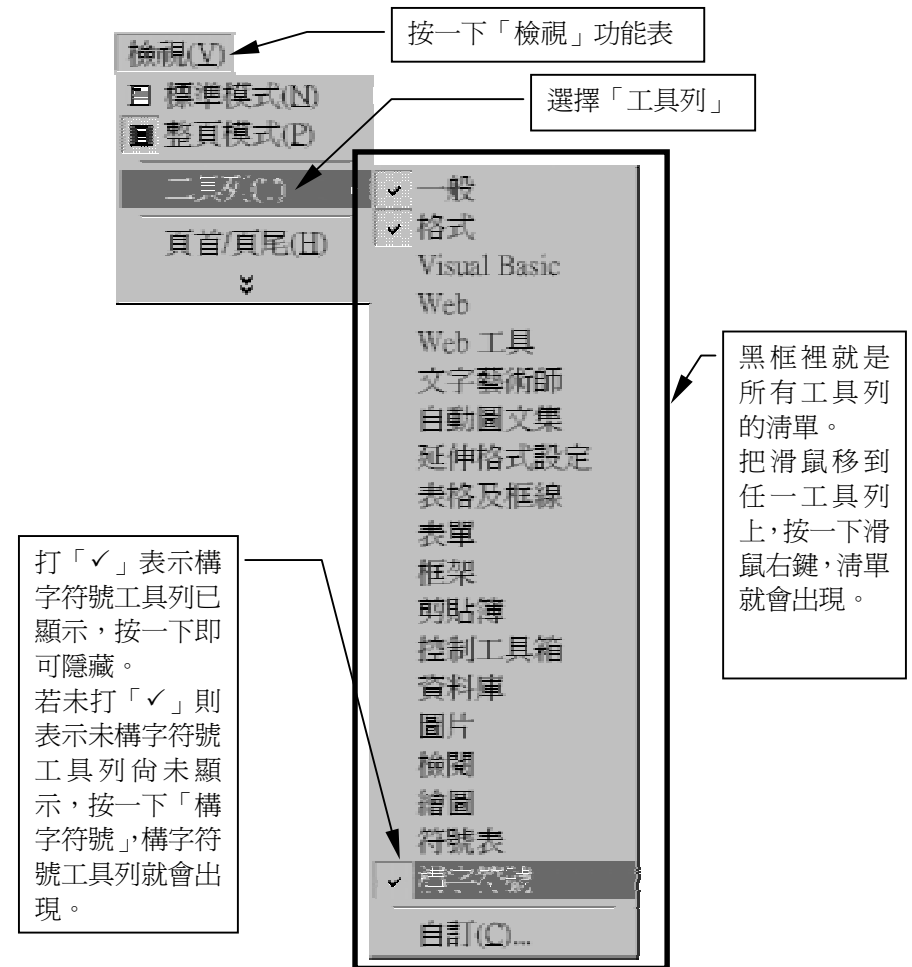


圖 2.3 顯示或隱藏構字符號工具列

3. 將滑鼠移到任一工具列上，按一下滑鼠右鍵，會跳出所有工具列的清單，前面有打「✓」的表示顯示的工具列。把滑鼠移到最下面的「構字符號」上，按一下，即可取消工具列顯示。

在 Word 處理缺字

在 Word 中缺字的輸入，是以**構字式**來表達。(構字式的使用方法，詳見附錄一)例如缺字「煊」的構字式為「火 Δ 宣」(表示「煊」字可由「火」及「宣」兩字橫連而得， Δ 表橫連)。若要閱讀及列印缺字，必須將**構字式**轉為**字形**或**圖片**。

在 Word 文件中，構字式可轉換成**字形**，但是此一功能在網頁及未安裝本系統的電腦上無法執行。這時就必須將構字式轉成**圖片**，以顯示正確的字形。

使用者輸入構字式的文件，稱為「原始文件」。透過缺字處理增益集，可將**構字式**轉成**字形**或**圖片**，並存於新增的文件中，這個新增的文件稱為「缺字預覽文件」。缺字預覽文件只供閱讀及列印，修改時必須回到原始文件進行。

• 構字式轉成字形

將構字式轉成字形的步驟如下：

例 2.1：輸入「煊」的構字式，並轉成字形

1. 輸入「火」
2. 按「構字符號工具列」上的「 Δ 」
3. 輸入「宣」
4. 儲存檔案，輸入檔名「A.doc」。
5. 按一下「工具」功能表，再按「構字式轉成字形☆(預覽)」(見圖 2.4)。

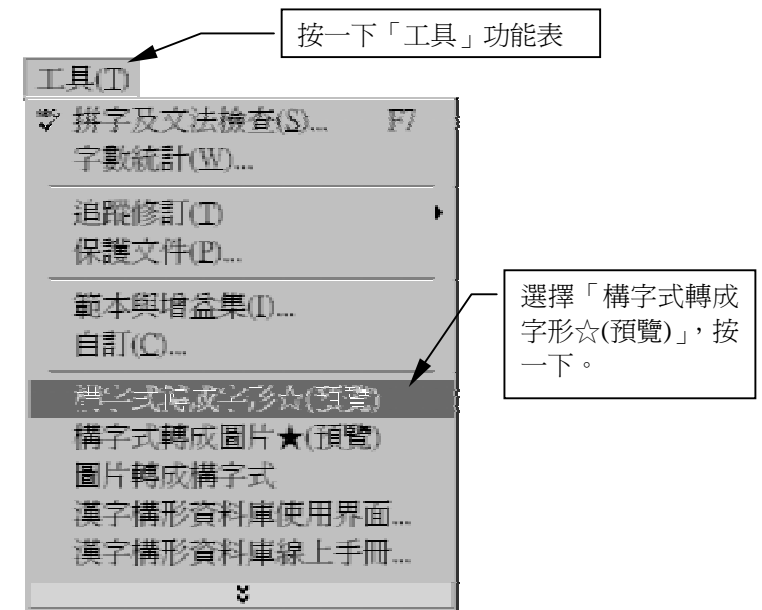


圖 2.4 工具功能表－「構字式轉成字形」

執行完畢，會新增一個「A☆.doc」文件，可供閱讀或列印。

- 構字式轉成圖片

將構字式轉成圖片的步驟如下：

例 2.2：承上例，將「煊」的構字式轉成圖片

1. 按一下功能表「視窗」、選「A.doc」。
2. 按一下「工具」功能表，之後再按「構字式轉成圖片★(預覽)」(見圖 2.5)。

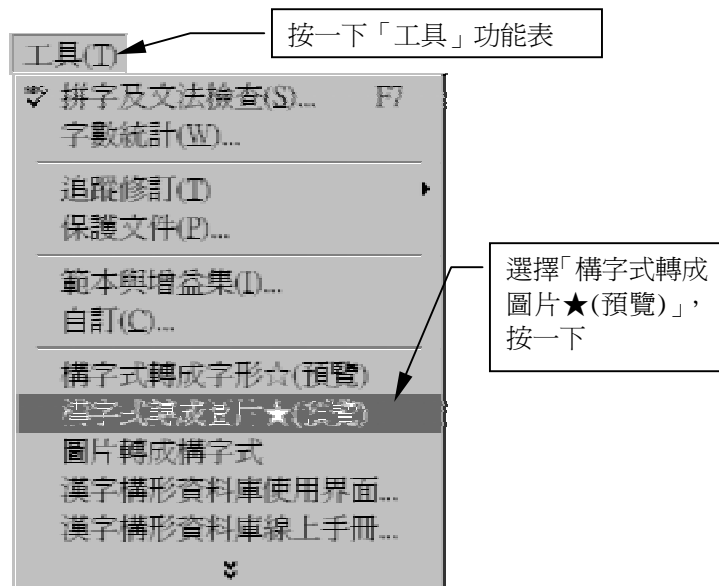


圖 2.5 工具功能表－「構字式轉成圖片」

執行完畢，會新增一個「A★.doc」文件，可供閱讀或列印。

說明：

原始文件在執行構字式轉換後，程式會自動新增缺字預覽檔案。由於缺字預覽可選擇轉成字形或圖片，預覽檔案的檔名，會分別加上「☆」及「★」以識區分。

現以「A.doc」為例，說明：

原 始 檔 案：A.doc
構字式轉成字形：A☆.doc
構字式轉成圖片：A★.doc

這三個檔案可以分別存取。例如，若需將含有缺字的文件轉送他人，可直接使用構字式轉成圖片的檔案，不需再附原始檔案。

- 檢視圖片中的構字式

構字式轉成圖片後，構字式會附加在圖片物件的替代文字裡，做為圖片轉回構字式的依據。

例 2.3：承上例，檢視圖片「煊」的構字式

1. 按一下功能表「視窗」、選「A★.doc」視窗。
2. 按一下「煊」，並複製「煊」的圖片。
3. 開啓一個新的 Word 檔案。
4. 貼上「煊」的圖片。
5. 按一下「煊」，選取圖片。
6. 按一下滑鼠右鍵，就會跳出一個功能表，選「圖片

格式」，再按一下（見第 15 頁，圖 2.6）。

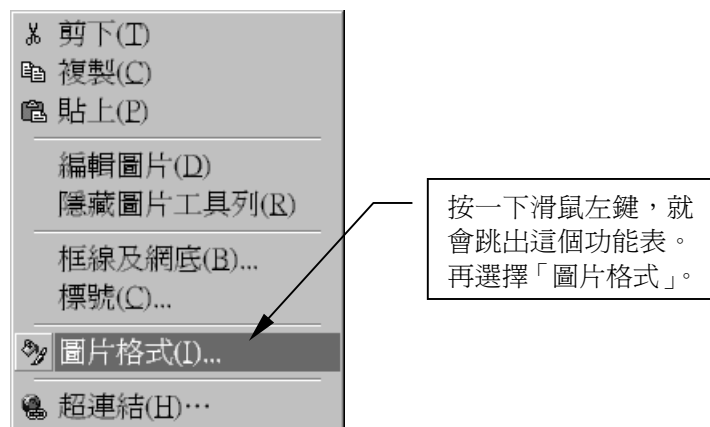


圖 2.6 滑鼠左鍵圖片功能表－「圖片格式」

7. 螢幕上出現圖片格式視窗，按一下「Web」（見圖 2.7）。

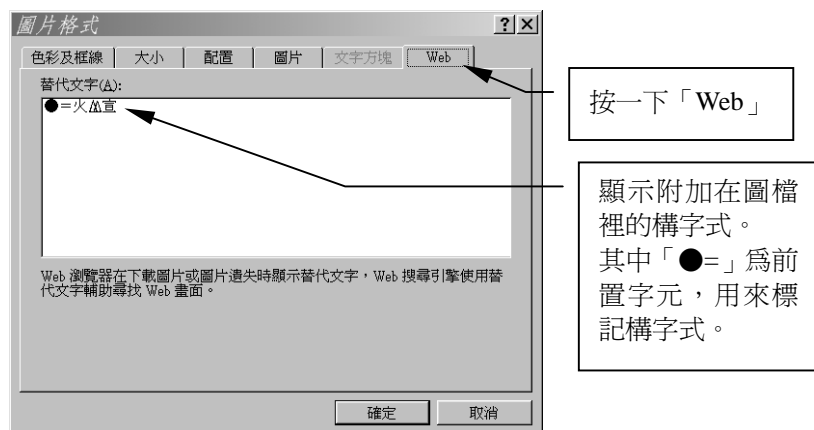


圖 2.7 「圖片格式」視窗

說明：

1. 並非每一個字形的圖片檔都可以轉成構字式，只有內含構字式的圖片，才能轉回構字式。
2. 「●=」為前置字元，用來標記構字式。

• 圖片轉成構字式

將圖片轉成構字式的步驟如下：

例 2.4：承例 2.3，將「烱」的構字式轉成圖片

按一下「工具」功能表，再按「圖片轉成構字式」。

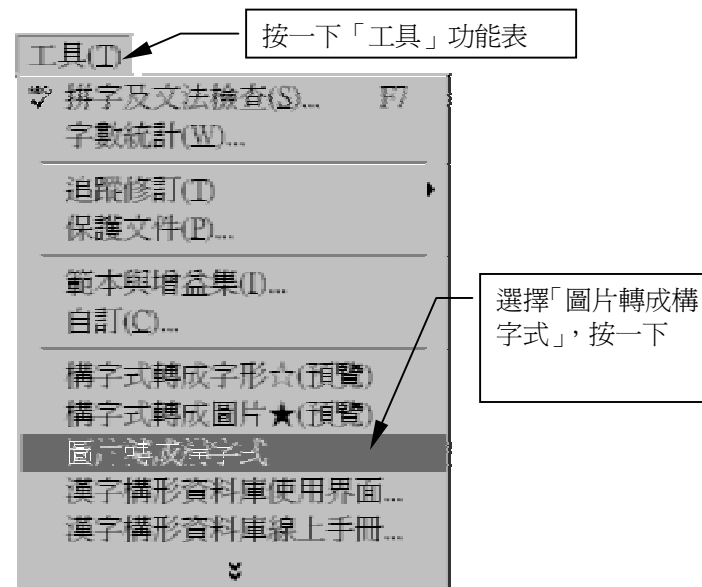


圖 2.8 工具功能表－「圖片轉成構字式」

執行完畢，「烺」的圖片會轉成構字式「火𤇀宜」。

• 解除文件保護

一般而言，含構字式的文件在轉成缺字預覽文件後，會加入「文件保護」以避免使用者任意修改。但若使用者想要直接編輯缺字預覽文件，就必須解除文件保護，才能編輯缺字預覽文件。

例 2.5：解除文件保護

承例 2.4

1. 按一下功能表「視窗」、選「A★.doc」視窗。
2. 按一下「工具」功能表，之後再按「圖片轉成構字式」。
3. 系統會跳出如圖 2.9 的對話框，按一下「確定」，關閉對話框。

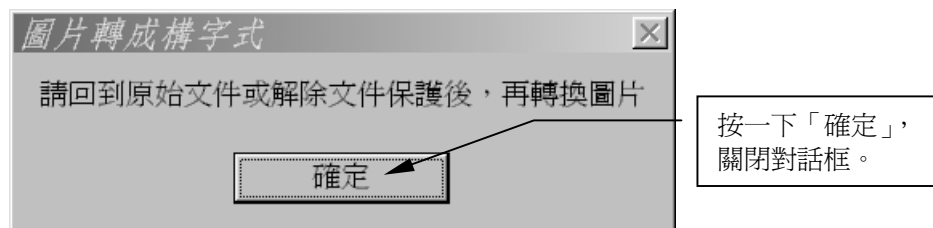


圖 2.9 「圖片轉成構字式」對話框

4. 按一下「工具」功能表，按一下「保護文件」，解

除文件保護（詳見第 18 頁，圖 2.10）。

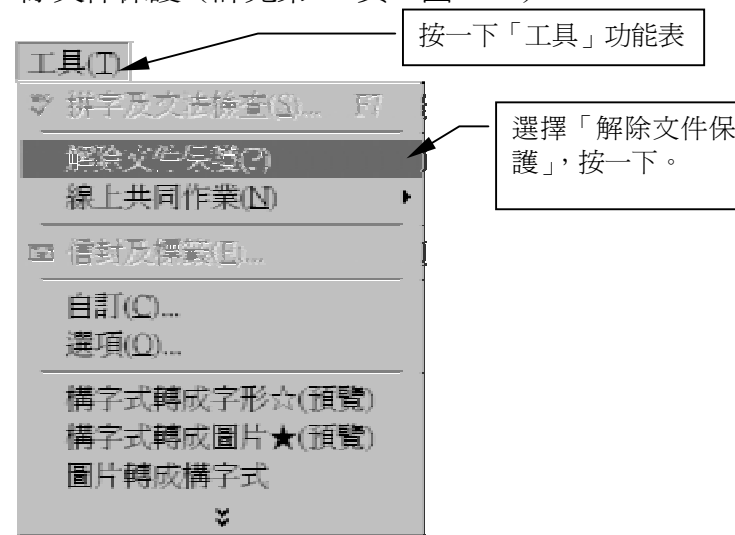


圖 2.10 解除文件保護

5. 按一下「工具」功能表，之後再按「圖片轉成構字式」。

即可看到「A★.doc」上的「烺」字的圖片，被轉換成構字式「火𤇀宜」。

注意：

除非特殊狀況，不建議使用解除文件保護此一功能，例如：使用者只擁有轉換過的缺字預覽檔案。缺字預覽文件都有其相對應的原始檔案，爲了確保原始檔和缺字預覽檔的對應正確，以及避免多次轉換後產生大量檔案的情形發生，**最好還是使用原始文件編輯或修改內容，再轉成各種缺字預覽文件。**

編輯文件內容

缺字預覽文件不能直接修改內容，必需切換回原始文件上才能編輯。

在原始文件「A.doc」新增缺字的程序如下：

例 2.6：新增缺字「凭」

承例 2.1

1. 按一下「視窗」功能表、然後選「A.doc」，回到原始文件的視窗。
2. 鍵入構字式「任厶几」（表示「凭」字可由「任」及「几」兩字直連而得，「厶」表直連，這個符號可由「構字符號工具列」中選取）。
3. 按功能表「工具」，之後再按「構字式轉成字形☆」。

此時，原始文件「A.doc」的內容如下：

火厶宣、任厶几

而缺字預覽文件「A☆.doc」的內容為：

煊、凭

注意：執行「構字式轉換」時會自動存檔，故此時不需再重覆例 2.1 的第 4 步驟（儲存檔案），直接執行「構字式轉換」即可。

構字式和一般文字

構字式和一般文字（非缺字）兩者可夾雜使用。

例 2.7：在一般文字中插入構字式

1. 在原始文件中輸入「怒髮衝冠任𡗗几欄處」。
2. 執行「構字式轉成字形☆」後，缺字預覽文件的內容為「怒髮衝冠凭欄處」。

構字輸入法

除了缺字以外，即使是一般文字，也可以構字式輸入。執行「構字式轉成字形☆」後，原始文件中的構字式會自動被一般文字取代。

例 2.8：以構字式輸入「蠶」

1. 在原始文件中輸入「偶遭蠱萬𡗗虫之讒」。
2. 執行「構字式轉成字形☆」，原始文件及缺字預覽文件的內容都是「偶遭蠱蠶之讒」。

在 Excel 處理缺字

使用 Excel 處理缺字，是用構字式來表達。

例 2.9：輸入「煊」的構字式，並轉成字形

1. 開啓一個新的 Excel 文件
2. 在儲存格 A1 中輸入構字式「火𡗗宣」
3. 儲存檔案，輸入檔名「A.xls」
4. 按一下「工具」功能表，再按「構字式轉成字形☆(預覽)」。

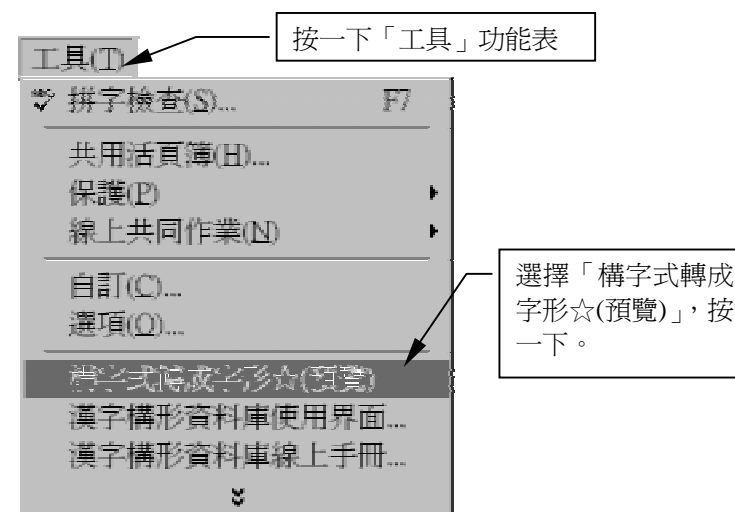


圖 2.11 Excel 工具功能表－「構字式轉成字型」

執行完畢，會新增一個「A☆.xls」文件，文件的儲存格 A1 中會顯示字形「煊」，可供閱讀或列印。

注意：

在 Excel 中，只能將構字式轉成字形，無法轉成圖片檔，也無法將圖片轉成構字式。

在 PowerPoint 處理缺字

使用 PowerPoint 處理缺字，是用構字式來表達。

例 2.10：輸入「煊」的構字式，並轉成字形

1. 開啓一個新的 Power Point 文件
2. 在文字方塊中輸入構字式「火 Δ 宣」
3. 儲存檔案，輸入檔名「A.ppt」
4. 按一下「工具」功能表，再按「構字式轉成字形☆(預覽)」。

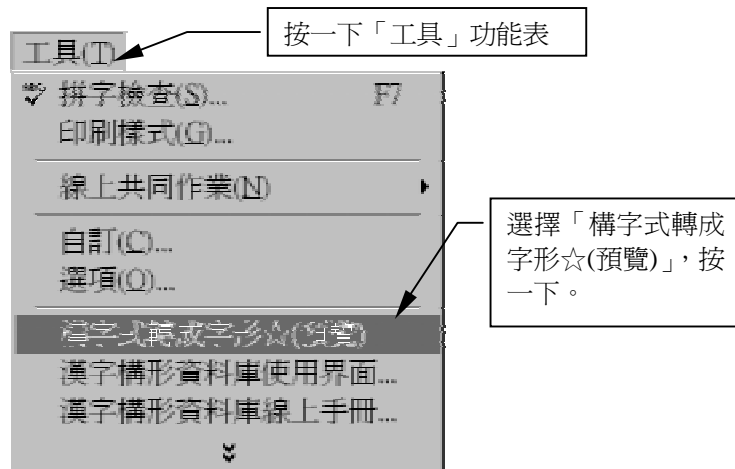


圖 2.12 PowerPoint 工具功能表－「構字式轉成字形」

執行完畢，會新增一個「A☆.ppt」文件，可供閱讀或列印。

注意：

在 PowerPoint 中，只能將構字式轉成字形，無法轉成圖片檔，也無法將圖片轉成構字式。

在 FrontPage 處理缺字

使用 FrontPage 處理缺字，是用構字式來表達。

例 2.11 編輯含缺字字形的網頁

1. 開啓 FrontPage，按一下「檔案」功能表，選「開啓 Web」，再按一下，會跳出一個「開啓 Web」視窗，選擇 Web 的目錄（預設為 My Web），按一下「開啓」，編輯一個新的 HTML 文件。
2. 輸入字串「怒髮衝冠任 Δ 几欄處」。
3. 按一下「工具」功能表，之後再按「構字式轉成字型★」。（見第 24 頁，圖 2.13）
4. 執行完畢，構字式「任 Δ 几」被轉換成圖片，網頁即顯示為「怒髮衝冠凭欄處」。
5. 按一下「檔案」功能表，選擇「儲存檔案」，再按一下，會跳出一個「另存新檔」的視窗，輸入檔名「A.htm」，再按一下「儲存」。按下「儲存」後，又會再跳出一個「儲存嵌入檔案」的視窗（見第 24 頁，圖 2.14），按「確定」，即可將「凭」的圖片檔以「cp3.gif」的檔名存到「image」目錄中。

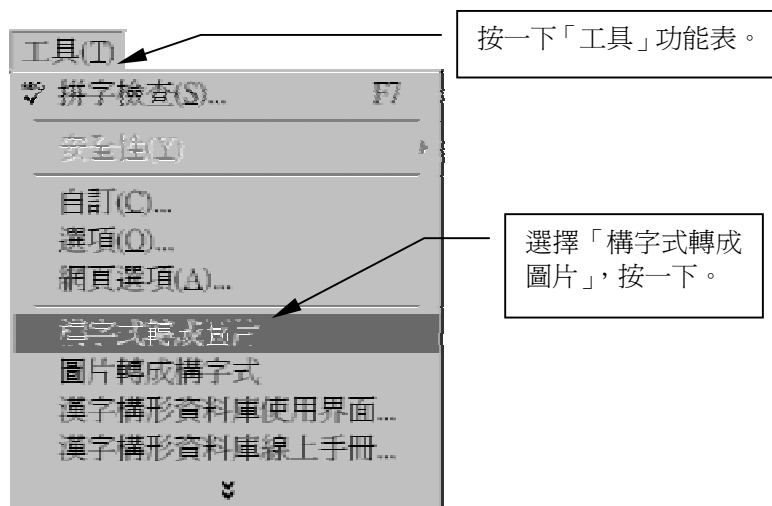


圖 2.13 Front Page 工具功能表－「構字式轉成圖片」

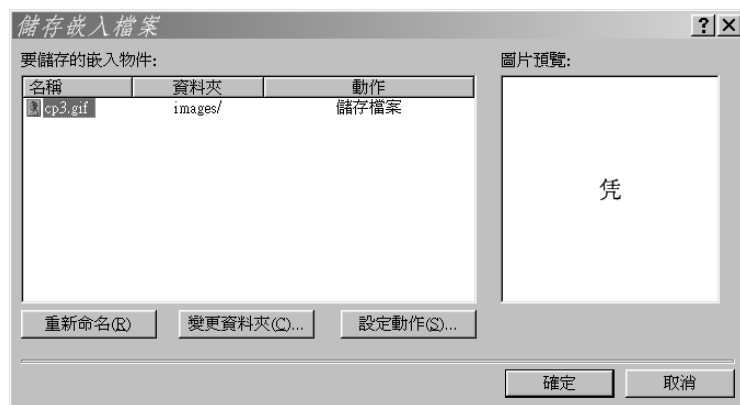


圖 2.14 儲存嵌入檔案

- 按一下檔案夾清單中的「A.htm」，選取，再按一下「檔案」功能表，選擇「用瀏覽器預覽」，再按一下。

即可瀏覽網頁內容。

注意：
如果不先將構字式轉成圖片，在網頁上只能看到構字式，無法看到字形。

例 2.12：用「几」或「壬」來查詢「凭」

在「字形孳乳」視窗上面的文字方塊內輸入「壬」或「几」，按「Enter」。視窗下面樹狀結構會顯示所有包含「壬」或「几」的字形（見圖 2.16）。

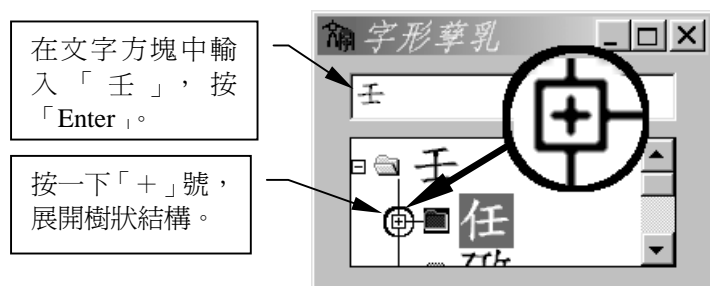


圖 2.16 展開樹狀結構

注意：

最好不要以「几」輸入。因為輸入「几」時，查詢出的字形個數比輸入「任」多，較不易找到「凭」。

例 2.13：展開或隱藏樹狀結構

承例 2.12

1. 按一下「任」前面的「+」號，樹狀結構會展開第二層，就可以找到「凭」（見圖 2.16）。
2. 按「-」號，即可隱藏「任」的樹狀結構（見第 29 頁，圖 2.17）。

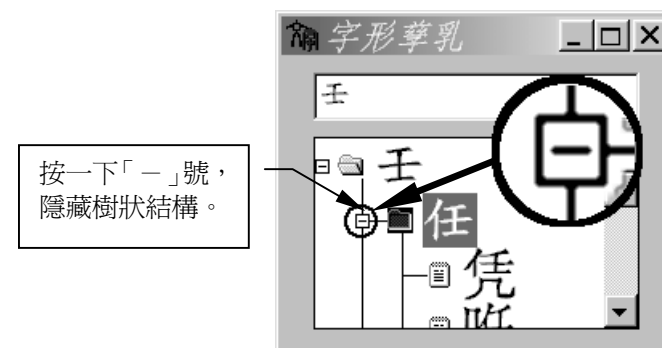


圖 2.17 隱藏樹狀結構

查詢字形（輸入多個部件）

文字方塊中可輸入多個部件，且不一定要依照順序輸入，亦不需輸入構字符號。

例 2.14：輸入「任几」來查詢「凭」

在「字形孳乳」視窗上面的文字方塊內輸入「任几」（見第 30 頁，圖 2.18），按「Enter」。視窗下面樹狀結構會顯示所有包含「几」和「任」的字形。

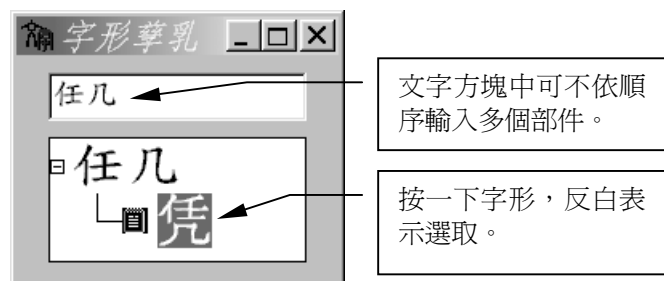


圖 2.18 多個部件查詢

複製構字式

選取「字形孳乳」視窗中的字形後，該字的構字式會自動複製到剪貼簿，其他的應用程式（例如 Word 等）就可以從剪貼簿中取得構字式。

例 2.15：取得「凭」的構字式

承例 2.14（見圖 2.18）

1. 按一下「凭」字，構字式會複製到剪貼簿中。
2. 切換回到 Word 視窗，按一下功能表「編輯」、「貼上」。就可以把構字式「任厶几」貼到 Word 文件中。

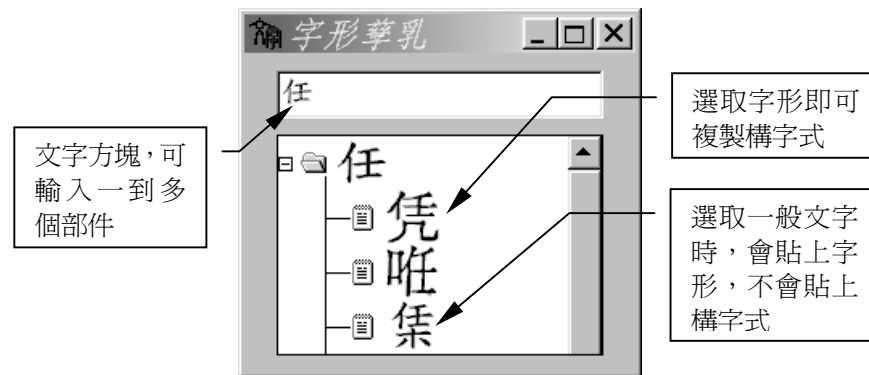


圖 2.19 複製構字式

注意：

如果選取的是一般文字時，會直接貼上字形，不會貼上構字式。

例 2.15 中步驟 1，若改選取「集」字。因「集」不是缺字，切換回 Word，貼上，會顯示「集」而非構字式「任厶木」。

查詢異體字

異體字表視窗，可以顯示選取字的異體字。

例 2.16：顯示「凭」的異體字

承例 2.14（見第 30 頁，圖 2.18）

按一下「凭」字，「凭」會反白表示已選取。即可顯示「凭」的異體字（見第 32 頁，圖 2.20）。

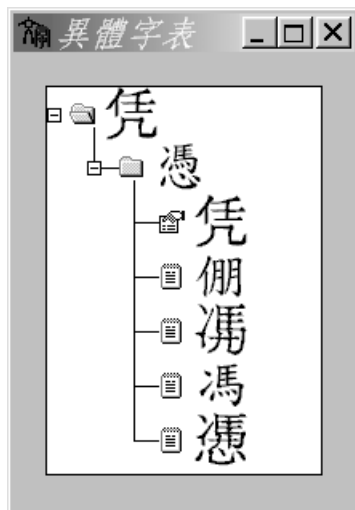


圖 2.20 顯示「凭」的異體字形

說明：

1. 在漢字構形資料庫使用界面中，按一下字選取之後，「異體字表」、「字形結構」及「字形演變」三個視窗中的資料，會隨選取字而改變。
2. 《漢語大字典》異體字表的編排，是將同一主體字統領的異體字集中在該主體字下編為一組。如圖 2.20 中「憑」為「凭」的主體字。

查詢字形源流演變

字形演變視窗，可以顯示選取字的字形源流演變。

例 2.17：顯示「凭」的字形演變

承例 2.14（見第 30 頁，圖 2.18）

按一下「凭」，「凭」字反白表示已選取。

「字形演變」視窗中即顯示「凭」字的小篆字形。（見第 30 頁，圖 2.18）

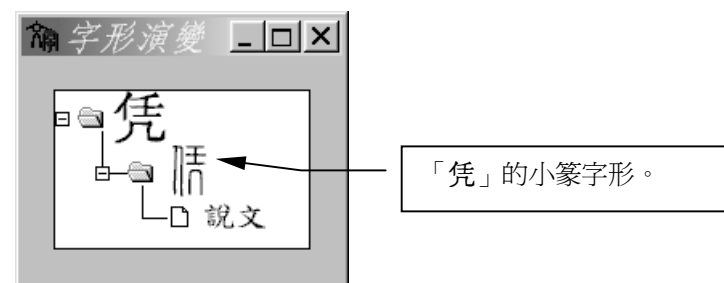


圖 2.21 「凭」的字形演變

說明：

若選取一字形後，字形演變視窗內只顯示所選取的字形，卻沒有列出對應的小篆字形，即表示《說文解字》未收錄此字形。

查詢字形結構

字形結構視窗，可以顯示選取字的字形結構。

例 2.18：查詢「凭」及其小篆「𠂔」的字形結構

承例 2.17（見第 33 頁，圖 2.21）

1. 先按一下「凭」，選取。顯示「凭」的字形結構。（見圖 2.22）
2. 再按一下小篆「𠂔」，選取。顯示「𠂔」的字形結構。（見圖 2.23）

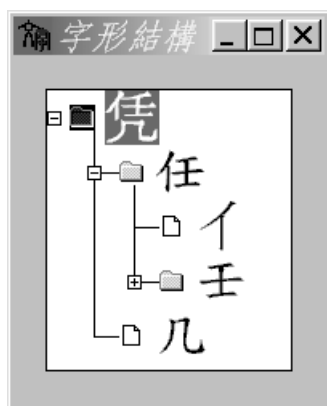


圖 2.22「凭」的字形結構

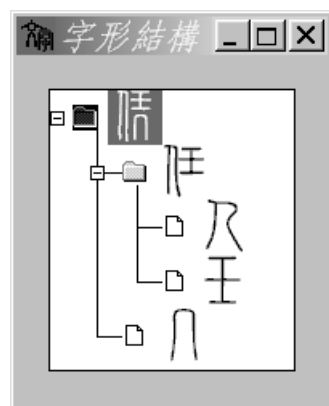


圖 2.23「𠂔」的字形結構

例 2.19：顯示「𠂔」的字形結構

1. 查詢「倍」的字形源流演變。
2. 按一下「倍」，選取（見第 35 頁，圖 2.24），字形結構視窗會顯示楷體「倍」的登錄結構，將這個結構完全展開（見第 35 頁，圖 2.25）。

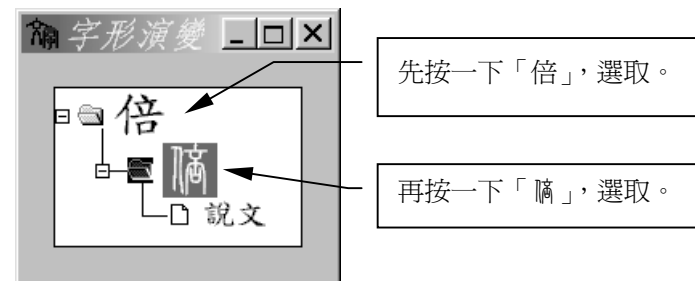
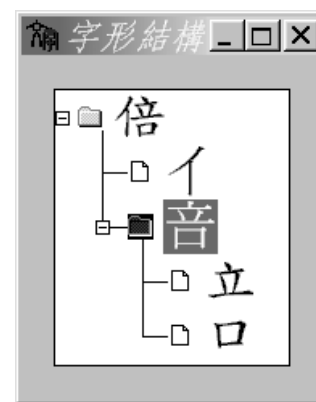
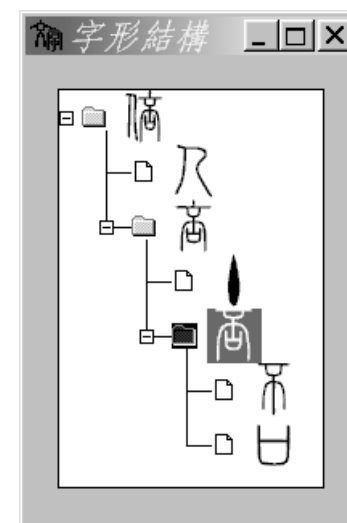


圖 2.24「倍」的字形演變

3. 再按一下小篆「𠂔」，選取（見圖 2.24），字形結構視窗會顯示小篆「𠂔」在《說文解字》的釋形，將這個結構完全展開（見圖 2.26）。



圖：2.25「倍」的登錄結構



圖：2.26「𠂔」在《說文》的釋形

三、網際網路的缺字處理

想要在網頁上顯示缺字，除了在 FrontPage 中將構字式轉成圖片之外，還可以採用在網頁上內含 Java Applet 的方式來處理。

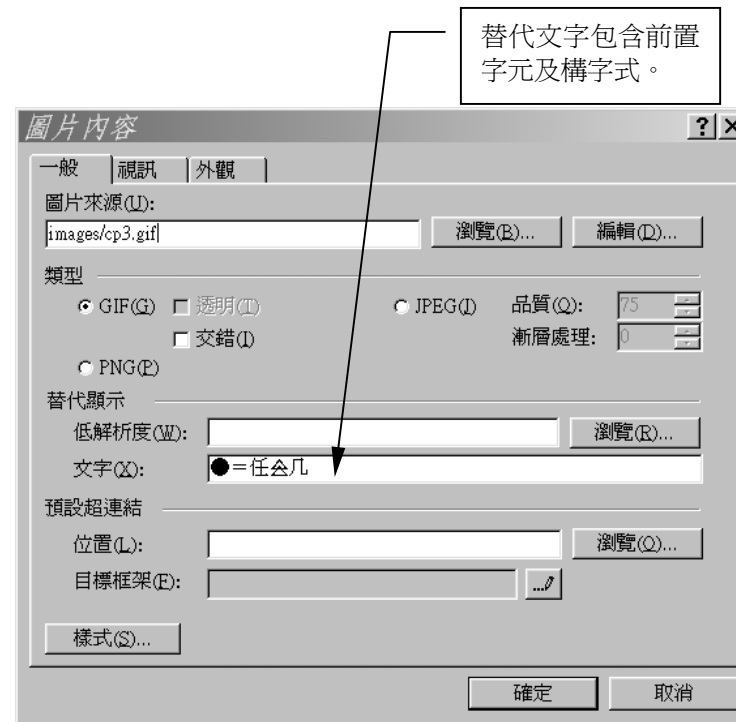
網頁的缺字表達

網頁的缺字係使用替代文字為構字式的圖片來表達。編輯網頁時可以利用本系統所提供的缺字處理增益集來轉換構字式，或是依照例 2.20 及 2.21 自行製作缺字圖片及標籤，即可表達網頁上的缺字。

例 2.20：FrontPage 中缺字圖片的內容

承例 2.11

1. 在「凭」的圖片上按一下滑鼠右鍵，選擇「圖片內容」，再按一下。開啓「圖片內容」視窗（見次頁，圖 2.27）。
2. 在「一般」標籤中，替代文字方塊中已填入構字式及前置字元。
3. 按一下「外觀」標籤，版面配置的對齊方式已改為「對齊文字頂端」（見次頁，圖 2.28）。



說明：

前置字元可為「●=」或「●=」（等號全、半形皆可）。

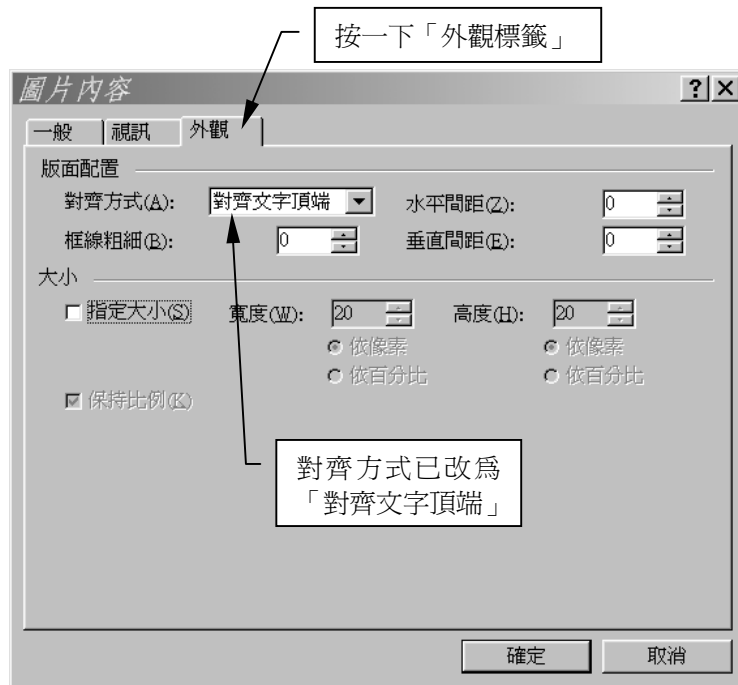
圖：2.27 缺字圖片的內容——一般

例 2.21：缺字圖片的標籤

承例 2.11

按一下「凭」的圖片，選擇。再按一下「HTML」標籤，即可顯示圖片「凭」的標籤，標籤的內容如下：

```
<img border="0" src=images/cp3.gif align="texttop"
width=20 height=20 Alt=●=任厶凡>
```

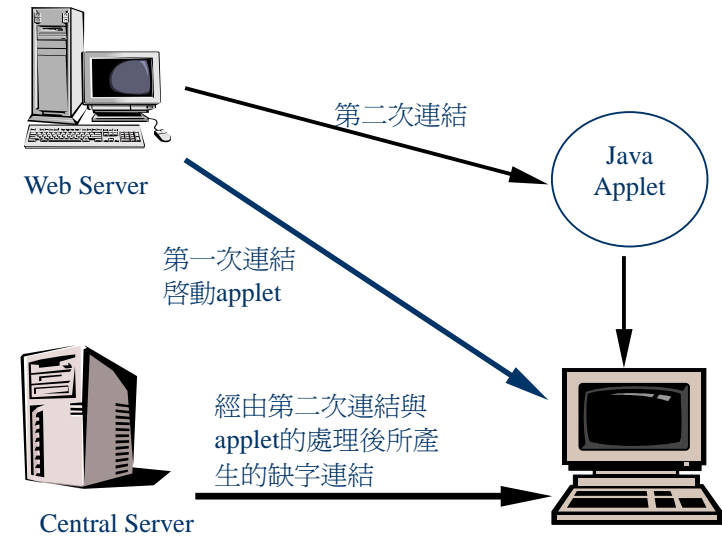


圖：2.28 缺字圖片的內容－外觀

使用 Java Applet 轉換網頁中的構字式

Java Applet 可將網頁中的構字式轉換成缺字圖片。瀏覽者下載網頁後，Java Applet 會要求網頁伺服器重新傳送網頁，並找出網頁中的構字式，傳給遠端

的伺服器轉換成圖片，再更新瀏覽器的視窗。運作架構如圖 2.29：



圖：2.29 Java Applet 運作架構圖

如何在網頁加入 Java Applet

先將 Java Applet 程式 (WordAppletn.jar) 置於 Web Server 上，再將含有構字式的網頁的 HTML 程式稍加修改如下：

將 Applet 程式置於<HTML>標題後，Applet 程式如下

所示：

```
<APPLET NAME=WordApplet ID=WordApplet  
    CODE=WordApplet.class  
    ARCHIVE="../WordAppletn.jar" WIDTH=150  
    HEIGHT=24 MAYSCRIPT>  
<PARAM NAME="LayerID"    VALUE="context">  
<PARAM NAME="FontSize"   VALUE="16">  
<PARAM NAME="FontColor"  VALUE=BLACK>  
<PARAM NAME="Visible"    VALUE="YES">  
</APPLET>
```

注意：

1. 網頁之 Charset 需設定為 Big5，方可執行。若未標示則預設為中文網頁，亦可執行。
2. NAME=WordApplet，表示該 Applet 的名稱為 WordApplet。
3. ID=WordApplet，表示該 Applet 的 ID 為 WordApplet。
4. CODE=WordApplet.class 表示該 Applet 要執行的程式為 WordApplet.class，而其程式被包於 WordAppletn.jar 中。
5. ARCHIVE="../WordAppletn.jar" 為該 Applet 所在的位置，路徑請寫相對路徑。例如：若該程式與網頁所在為同一目錄下則可寫作 ARCHIVE="WordAppletn.jar"。若該程式於上一層目錄中，則可寫作 ARCHIVE="../WordAppletn.jar"。
6. WIDTH=150 HEIGHT=24 則為該 Applet 呈現於網頁中的大小，若要完全消失請將設定設成 WIDTH=0 HEIGHT=0，MAYSCRIPT 則為必須填寫，其餘皆不必更改。

參數部分：

1. <PARAM NAME="LayerID" VALUE="context">
表示需要轉換缺字的部分，其物件的名稱或 ID 為 context，此名稱或 ID 可是網頁的需要而修改。此參數為必須填寫。
2. <PARAM NAME="FontSize" VALUE="16">
表示轉換後的缺字字形大小，此參數為選擇性填寫，若未填寫，其與預設值為 12。
3. <PARAM NAME="FontColor" VALUE=BLACK>
表示轉換後的缺字字形顏色，此參數為選擇性填寫，若未填寫，其與預設值為黑色。
4. <PARAM NAME="Visible" VALUE="YES">
表示該 Applet 是否呈現於網頁中的選項，此參數為選擇性填寫，若未填寫，其與預設值為 YES，若不要將 Applet 呈現於網頁中，可將參數設成 NO 即可。
5. 最後的步驟就是將網頁的內文以 <DIV> 夾在其中，如下：
<DIV ID="context" STYLE="visibility:show;position:absolute">
.....
.....內文.....
.....
</DIV>
請注意 ID="context"，ID 的值要與第一項參數 VALUE="context" 相同，這樣 Applet 才知道要轉換的區域為 ID="context"的 <DIV> TAG 內。

上述步驟設定好後，缺字轉換程式即可運作。

說明：WordAppletn.jar 檔案可於漢字構形資料庫安裝目錄中取得。

使用 Microsoft Office 處理網頁缺字

網頁中的缺字圖片可複製到 Microsoft Word 或 FrontPage 中，但無法直接複製到 Excel 及 PowerPoint 中，必須先透過 Word 或 FrontPage 轉換成構字式後，才能使用。

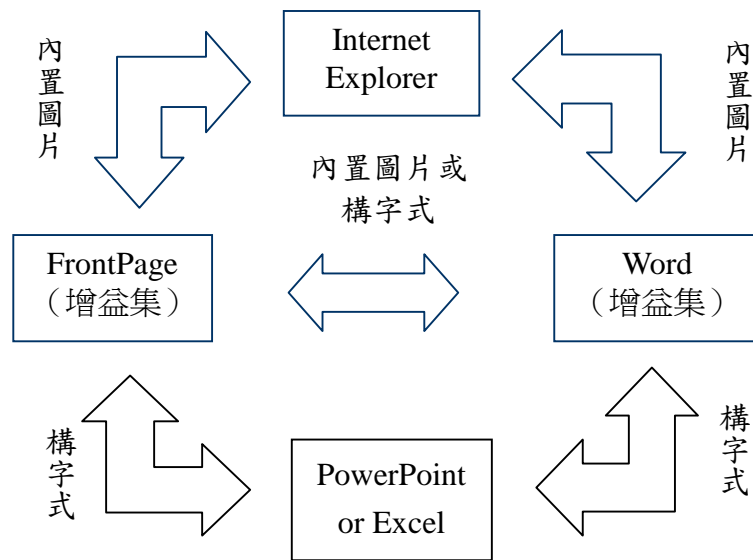


圖 2.30 網頁的缺字處理流程

在網際網路上查詢缺字與異體字

除了使用漢字構形資料庫使用界面之外，網際網路上也可以查詢缺字與異體字，相關功能請參考文獻處理實驗室網站：

<http://gate.sinica.edu.tw/~cdp>

• 缺字查詢

任几	送出查詢	a
凭	任会几	任会几
共【1】筆 搜尋關鍵字：【任几】 關鍵字字根：【几一士】		

圖 2.31 缺字查詢網站畫面

- 異體字查詢

漢語大字典異體字表 (Unicode 3.2)

[Full](#)

(包含 CJK 認同表
義文字基本集，
Extension A，
Extension B)

[Small](#)

(包含 CJK 認同表義文字
基本集，
Extension A，不含
Extension B)

● 3400-4DFF : CJK 認同表義文字 Extension A
 ● 4E00-9FFF : CJK 認同表義文字 基本集
 ● 20000-2A6DF: CJK 認同表義文字 Extension B

異體字查詢

圖 2.32 異體字查詢網站首頁

search: [回首頁](#)

組別	型別	Unicode	字型
11668	4	U+04CF6	騏
11668	3	U+096DE	雞
11668	1	U+09DC4	鷄
11668	2	U+09E21	鸡

圖 2.33 網路異體字查詢結果

第三章 構字式的自動校正

一般而言，執行構字式轉換時不會改變原始文件的內容，但是遇到使用者輸入的構字式和系統登錄不同的情況時，系統會修改原始文件中的構字式（詳見附錄一）。

一、一般文字的構字式

若構字式對應到一般文字，系統會將原始文件中的構字式取代為一般文字。

如，輸入「馮厶心」，執行「構字式轉換」時，原始文件中的構字式會由「憑」取代。

二、同一字形的不同構字式

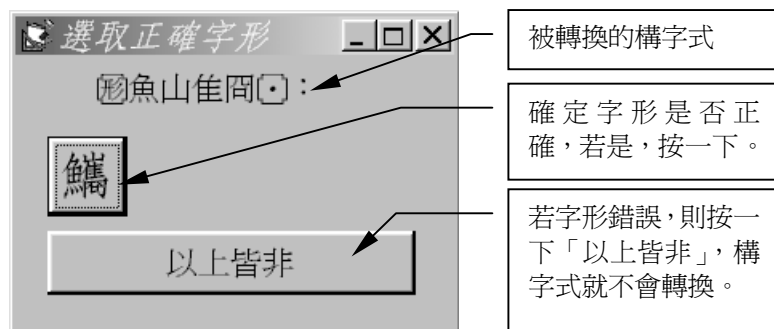
為兼顧使用上的方便，一個字因拆分方式不同，可以有多種不同的構字式。執行「構字式轉換」後，原始文件中的構字式會由最精簡的表達方式取代。

例 3.1：以不同的構字式輸入「鱻」

輸入以下三種構字式

- 𩺰魚山佳冏
- 𩺰魚崔冏
- 魚𩺰

1. 執行「構字式轉成字形☆（預覽）」功能。
2. 系統跳出「選取正確字形」對話方塊，讓使用者選擇構字式「𩺰魚山佳冏」的正確字形。
3. 按一下「鱻」字形的按鈕（見圖 3.1）。



圖：3.1 選取正確字形對話盒（一）

4. 系統再次跳出「選取正確字形」對話方塊，讓使用者選擇構字式「𩺰魚崔冏」的正確字形。
5. 按一下「鱻」字形的按鈕。

6. 在「構字式轉成字形☆（預覽）」文件中，三個構字式都被轉換成「鱻」。
7. 切換到原始文件，原始文件中的構字式已被代換。

由於「魚𩺰」最為精簡，原始文件中的構字式「𩺰魚山佳冏」或「𩺰魚崔冏」，都由「魚𩺰」取代。

例 3.2：以不同的構字式輸入「𩺰」字

- 𩺰魚魚魚
- 𩺰魚𩺰魚
- 𩺰魚𩺰魚
- 𩺰魚

由於「𩺰魚」最為精簡，在執行「構字式轉換」後，原始文件中其他的構字式，都會由「𩺰魚」取代。

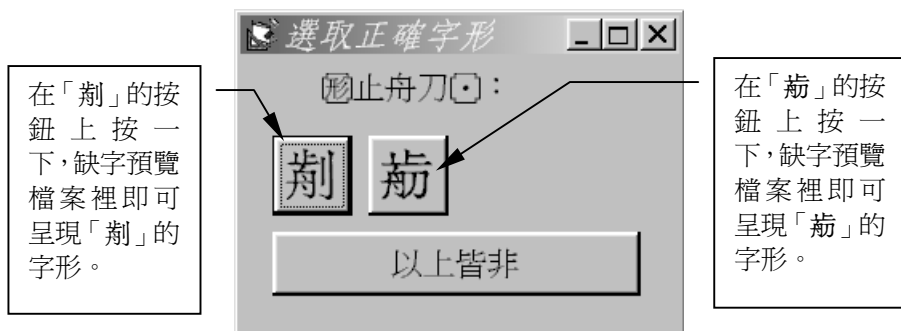
三、字形相似的構字式

有一些字形非常相似，如「𦵏」和「𦵐」。當輸入的構字式不夠明確，以致系統無法確認。系統就會列出所有相似字形，讓使用者選取。選取後，原始文件中的構字式會由較明確的構字式取代。

例 3.3：輸入「𦵏」（構字式為「止𠂔𠂔」）

1. 輸入構字式「𠂔止舟刀𠂔」。
2. 執行「構字式轉換」。
3. 螢幕上出現「𦵏」、「𦵐」二個字形。（見圖 3.2）
4. 選取「𦵏」。

原始文件中的構字式「𠂔止舟刀𠂔」，會由「止𠂔𠂔」取代。



圖：3.2 選取正確字形對話盒（二）

四、包含異寫部件的構字式

部首類的部件，如「氵」、「灬」、「刂」等，可以改用「水」、「火」、「刀」等異寫部件輸入。

例 3.4：輸入「洊」字

1. 輸入「水𠂔冬」
2. 執行「構字式轉換」
3. 原始文件中的構字式會由「氵𠂔冬」取代。

其他字舉例如表 3-1 所示：

表 3-1 異寫部件舉例

部首	異寫	例字	輸入的構字式	取代的構字式
火	灬	杰	不𠂔火	不𠂔灬
刀	刂	刘	文𠂔刀	文𠂔刂
手	扌	材	手𠂔才	扌𠂔才
心	小	悽	心𠂔求	小𠂔求

第四章 缺字字型

相對於一般文字，使用 Word 更換缺字字型或複製缺字時，應注意以下事項。

一、支援單一字型

目前缺字字型只有一種，在原始文件變更字型時，缺字的字型不會改變。

例 4.1：變更「陶喆」的字型

執行程序為：

1. 在原始文件中以新細明體輸入「陶〇〇吉」
2. 改以標楷體顯示「陶〇〇吉」
3. 執行「構字式轉換」
4. 缺字預覽文件中顯示「陶喆」

二、複製缺字構字式

從 Word 複製缺字時，應複製原始文件中的構字式。

例 4.2：複製缺字「喆」的構字式到記事本：

1. 在原始文件中輸入「〇〇吉」，複製構字式「〇〇吉」。
2. 按一下「開始」，再按一下「附屬應用程式」，選「記事本」。
3. 按一下「記事本」的功能表「編輯」、「貼上」，即貼上構字式「〇〇吉」。

不要直接複製缺字預覽文件中的字形，以免文件內容發生錯誤。以下兩例為錯誤的示範：

錯誤的程序例一

例 4.3：直接複製缺字「喆」到記事本

1. 複製缺字預覽文件中的字形「喆」。
2. 按一下「開始」，再按一下「附屬應用程式」，選「記事本」。
3. 按一下「編輯」，再按一下「貼上」。貼入的字形是「𠄎」。

即使在 Word 文件中，也儘量避免直接複製缺字字形，最好以構字式來表達缺字。

錯誤的程序例二

例 4.4：直接選取「喆」來變更字形：

1. 複製缺字預覽檔案中的缺字「喆」。
2. 開啓一個新的 Word 文件，貼上「喆」。
3. 變更「喆」的字型為標楷體或其他字型，則顯示「𠄎」。

到步驟 2 時，貼上的字型是正確的，但是進入步驟 3 變換字型後，就會產生錯誤。

三、古文字字型

因應古文字電腦化的需求，缺字預覽目前可支援「北師大小篆」字型。

例如，《說文》釋「玩」字如下：

𠄎，弄也，从玉元聲；𠄎，玩或从貝。

其中「𠄎」和「𠄎」為小篆，對應的楷化字分別為「玩」和「貶」。但「貶」為缺字，必須用構字式「貝△元」來表達。字形間的關係如下表：

表 4-1 小篆、楷化字及構字式的對應表

小篆	楷化字	構字式
玨	玩	王厶元
𠂔	𠂔	貝厶元

操作方式如下：

例 4.5：顯示「玨，弄也，从玉元聲；𠂔，玩或从貝。」

注意：

「玨」為「玩」的小篆。如果小篆字是一般文字（非缺字），直接輸入即可，不需使用構字式。

1. 輸入以下文字

玩，弄也，〇〇人玉元聲；貝厶元，玩或〇〇人貝。

2. 分別選取「玩」和「貝厶元」，字型改為「北師大小篆」。顯示如下：

玨，弄也，〇〇人玉元聲；貝厶元，玩或〇〇人貝。

3. 執行「構字式轉換」，顯示如下：

玨，弄也，从玉元聲；𠂔，玩或从貝。

第五章 實例輸入

本章的目的在提供實際應用的完整範例，做為練習的參考。希望透過實際文件輸入的過程，為使用者建立操作系統的概念。

本範例採自《說文解字注》一書中釋「教」字的段落，模擬一個使用者想將該段落輸入成一篇 Word 文件可能遇到的問題和解決的方法。

一、文件輸入須符合的條件

原始文件呈現的樣式是版刻書的書影（見次頁，圖 5.1），輸入時希望能儘量保留較多原始文件的資訊。如：呈現小篆和重文（異體字）字形，使用刻版的字形，而不使用今體的替代字等。

經分析後，歸納四項要求如下：

1. 直接輸入缺字，如「攴」。
2. 依原本呈現的字形輸入，不使用替代字，如「从」不以「從」代。
3. 輸入小篆字形及重文，如「𠂔」及「𠂔」。
4. 小字的夾註部分以符號☆及★包夾，如☆從攴從爻★。

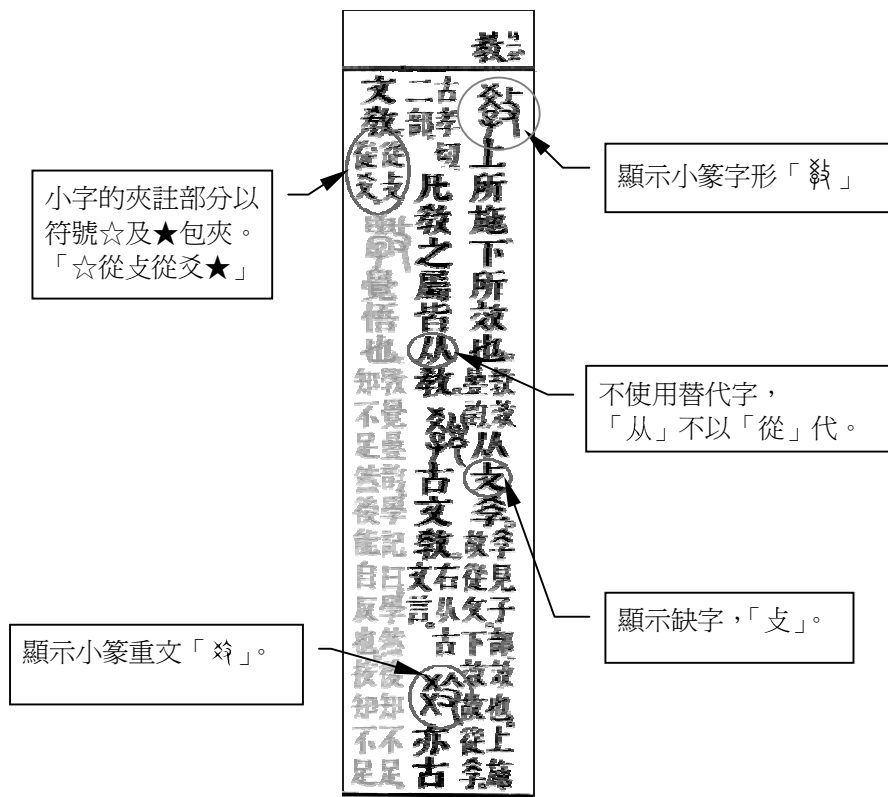


圖 5.1：《說文解字注》釋「教」

二、缺字的輸入

使用漢字構形資料庫可以解決缺字（含小篆）的輸入，以下以實例來說明操作程序：

安裝漢字構形資料庫

安裝的細節詳見第一章，此處不再贅述。

使用 Word 來輸入《說文》釋「教」的段落

安裝漢字構形資料庫後，即可開啓 Word，在文件中輸入缺字。

在輸入時，需要解決的缺字問題如下：

1. 小篆字形「攴」的顯示。
2. 輸入「彊」、「韻」、「从」、「攴」、「季」、「攴」等缺字。
3. 輸入小篆重文「攴」、「攴」。

以下依序說明輸入的方式。

- 輸入小篆字形「𠄎」

教上所施下所效也。☆教效

選取「教」字，將字型換成「北師大小篆」，即可顯示：

𠄎上所施下所效也。☆教效

說明：「☆」、「★」兩個符號，可以使用 Word 中插入特殊字元的功能來輸入。

- 輸入缺字「𠄎」

缺字須使用構字式表達，再轉換成字形。構字式的詳細說明詳見附錄一。構字式可自行輸入，也可以使用漢字構形資料庫使用介面來複製。漢字構形資料庫使用介面的使用詳見第二章第二節。

使用漢字構形資料庫查詢字形時，僅需輸入部分的字形即可。以「𠄎」字為例，使用者可能只認識上面有一個「晶」，其他部分無法輸入。但是使用者可以只輸入「晶」，就可以查詢到「𠄎」，並取得構字式。

使用「晶」來取得「𠄎」的構字式的程序如下：

1. 開啓漢字構形資料庫使用界面

按一下功能表「工具」，選「漢字構形資料庫使用界面」，開啓漢字構形資料庫使用者界面的視窗。（見圖 5.2）

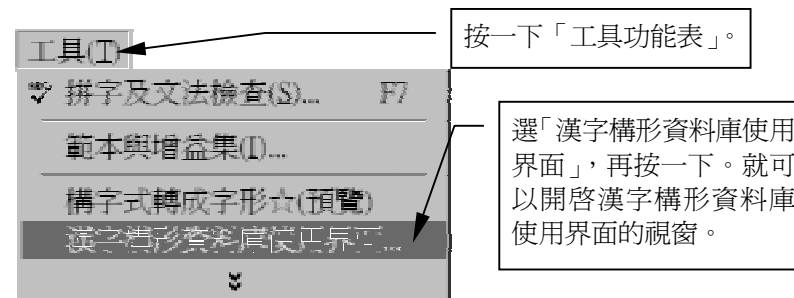


圖 5.2：開啓漢字構形資料庫使用界面

2. 輸入「晶」查詢字形

在字形孳乳視窗上方的文字方塊中輸入「晶」，輸入完畢後，按一下「Enter」，查詢的結果會顯示所有包含「晶」的字形。（見次頁，圖 5.3）

3. 複製「𠄎」的構字式

找到「𠄎」，按一下字形（見圖 5.3），被選取字的構字式就會自動複製到剪貼簿。

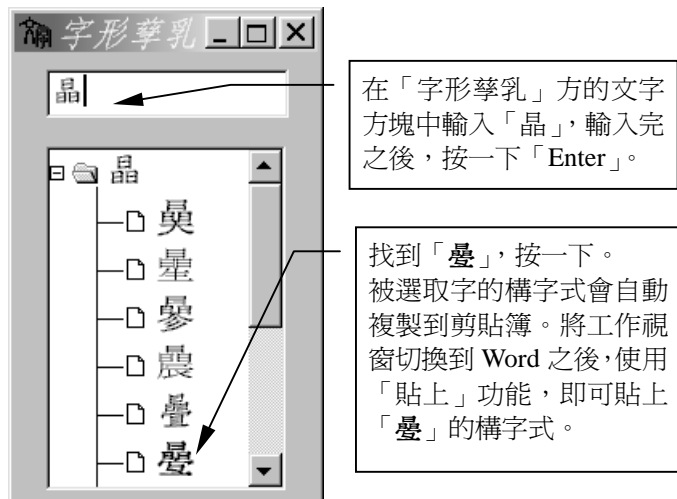


圖 5.3：查詢「𪛝」的字形並複製構字式

4. 到 Word 裡貼上「𪛝」的構字式
切換到 Word 視窗，按一下滑鼠右鍵，選擇「貼上」。
即可顯示：

𪛝上所施下所效也。☆教效晶𪛝𪛝𪛝

5. 將構字式「晶𪛝𪛝𪛝」轉成字形「𪛝」
 - a. 儲存檔案，將檔名設為「教.doc」。
 - b. 按一下「工具」功能表，選擇「構字式轉成字形 ☆ (預覽)」。(見次頁，圖 5.4)

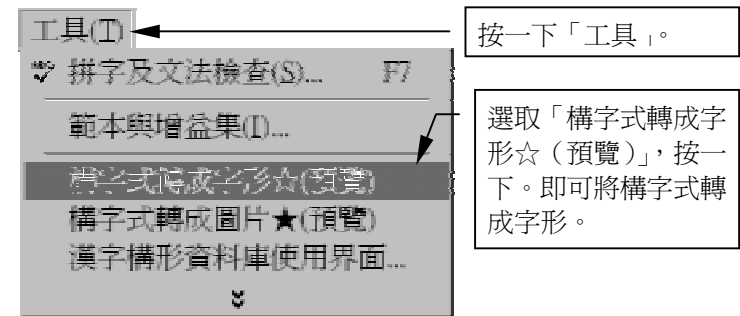


圖 5.4：將構字式轉成字形

- c. 構字式轉換完成，並新增一個文件「教☆.doc」。

在文件「教☆.doc」中即可顯示正確的字形：

𪛝上所施下所效也。☆教效𪛝

說明：構字式轉成字形可在文件全部輸入完之後一次將構字式全部轉換，不需要每輸入一個構字式就轉換一次。

• 輸入缺字「韵」

漢字構形資料庫也可以查詢異體字。若使用者已知「韵」是「韻」的異體字，即可輸入 Big-5 有碼的字形「韻」來查詢異體字「韵」，並取得異體字的構字式。

1. 輸入「韻」查詢異體字形

在字形孳乳視窗上方的文字方塊中輸入「韻」，輸入完畢後，按一下「Enter」。字形孳乳下面的方框就會顯示所有含「韻」的字形，用滑鼠在「韻」的字形上按一下，選取「韻」的字形。（見圖 5.5）

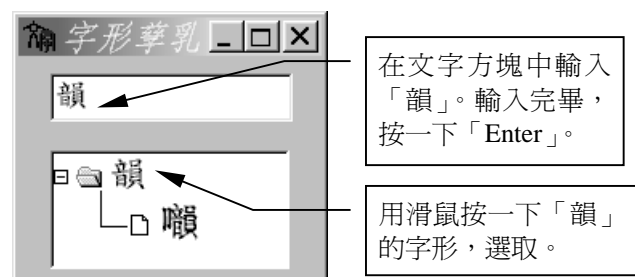


圖 5.5：查詢「韻」的異體字

選取字的異體字會顯示在「異體字表」的視窗裡，只要選取「韻」，「韻」的異體字就會全部列在「異體字表」的視窗裡。（見次頁，圖 5.6）

2. 複製「韵」的構字式

找到「韵」，按一下字形（見圖 5.6），被選取字的構字式就會自動複製到剪貼簿。

3. 到 Word 裡貼上「韵」的構字式

切換到 Word 視窗，按一下滑鼠右鍵，選擇「貼上」。即可顯示：

𠂔上所施下所效也。☆教效晶𠂔𠂔互
𠂔音勺? 𠂔★

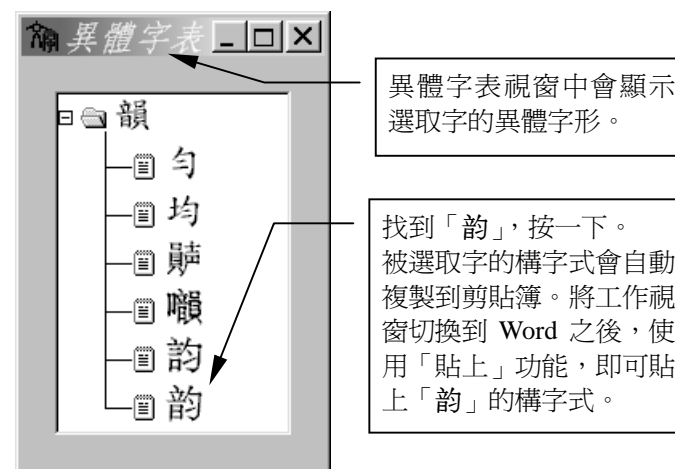


圖 5.6：「韵」的異體字查詢結果

• 輸入缺字「从」

輸入構字式時，並不一定要查詢資料庫。使用者可以自行拆分，直接輸入各種構字式。

構字式中有一組「方便符號」，專門用來表達由一個部件重覆組合成的字形。例如：「林」、「森」等字形。「从」字由兩個「人」組成，故可以構字式「人人」

來代表「从」字。

構字符號可以直接輸入，或是使用構字符號工具列來協助輸入。構字符號的輸入法，詳見附錄二。構字符號工具列的使用詳見第九頁。

1. 開啓構字符號工具列

按一下「檢視」功能表，選「工具列」，會跳出所有工具列的清單。將滑鼠移到最下面未打「✓」的「構字符號」上，按一下，即可顯示。(見第 10 頁，圖 2.3)

2. 選取構字符號「∞」

按一下構字符號工具列上的符號「∞」。(見圖 5.7)

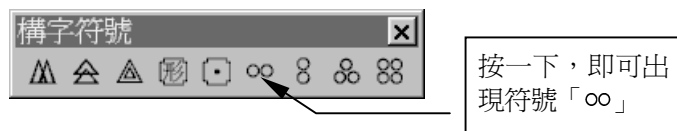


圖 5.7：選取構字符號「∞」

3. 輸入重複字形「人」

將字形重複出現的部分輸入在方便符號「∞」之後，構字式「∞人」即可代表字形「从」。

𠂔上所施下所效也。☆教效晶𠂔𠂔𠂔
𠂔音勺? □★∞人

• 輸入缺字「攴」

除了查詢字形的衍生字，漢字構形資料庫還可以取出部分的字形。例如「攴」為字形「敲」的一部分，即可以利用漢字構形資料庫使用界面來輸入「攴」。

1. 輸入「敲」查詢字形

在字形孳乳視窗上方的文字方塊中輸入「敲」，輸入完畢後，按一下「Enter」。字形孳乳下面的方框就會顯示所有含「敲」的字形，用滑鼠在「敲」的字形上按一下，選取「敲」的字形。(見圖 5.8)

選取字的字形結構會顯示在「字形結構」的視窗裡，「敲」的字形結構視窗裡會顯示敲是由「高」和「攴」組成。(見次頁，圖 5.9)

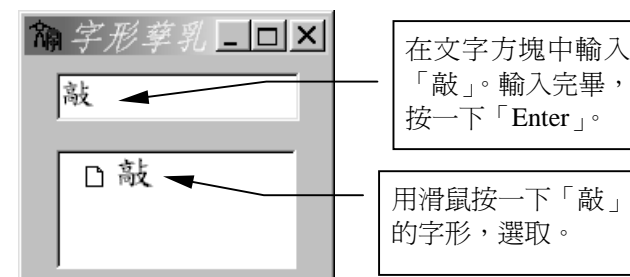


圖 5.8：查詢「敲」

2. 複製「攴」的字形

找到「攴」，按一下字形（見次頁，圖 5.9），「攴」

的字形就會自動複製到剪貼簿。

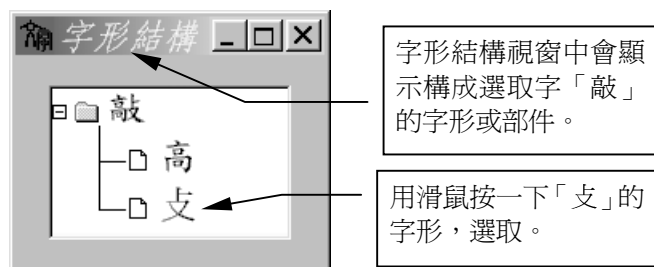


圖 5.9：複製「支」的字形

3. 到 Word 裡貼上「支」的構字式
切換到 Word 視窗，按一下滑鼠右鍵，選擇「貼上」。
即可顯示：

𠂔上所施下所效也。☆教效晶𠂔𠂔𠂔
𠂔音勺？□★∞人支爻𠂔子。☆爻𠂔子
見子部。效也。上施故從

• 輸入缺字「𠂔」

除了使用方便符號輸入構字式之外，一般的構字式也可以直接輸入。例如「𠂔」，就是上「爻」下「子」所組合成，使用者可以不必每次都查詢資料庫後再將構字式複製到 Word 文件，直接使用鍵盤輸入「爻𠂔子」即可。

1. 輸入爻

2. 選取構字符號「𠂔」

按一下構字符號工具列上的符號「𠂔」。(見圖 5.10)

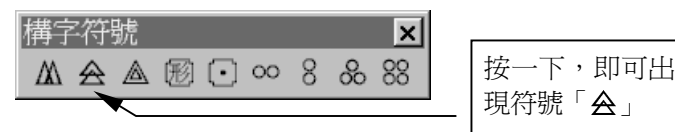


圖 5.10：選取構字符號「𠂔」

3. 輸入「子」和之後的文字
即可顯示：

𠂔上所施下所效也。☆教效晶𠂔𠂔𠂔
𠂔音勺？□★∞人支爻𠂔子。☆爻𠂔子
見子部。效也。上施故從

• 輸入缺字「𠂔」

「𠂔」是「支」的異體字，使用者可以切換到「漢字構形資料庫使用界面」的視窗，選取「支」的異體字形（見次頁，圖 5.11），再貼到 Word 文件即可。

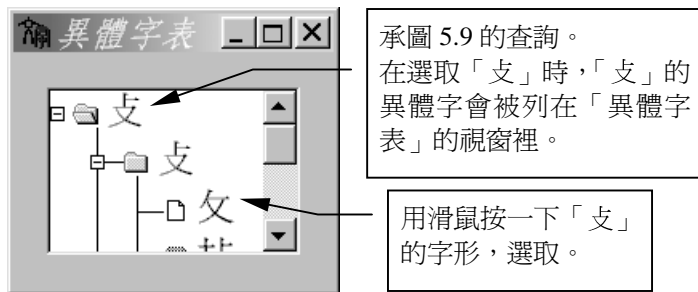


圖 5.11：複製字形「攴」

除了上述方式之外，由於漢字構形資料庫的外字集有收錄「攴」，並編好倉頡和注音輸入法，使用者也可以使用鍵盤輸入字形「攴」。輸入碼取碼方式的簡要說明和常用部件的輸入法，請參閱附錄四。

使用鍵盤輸入「攴」和其他文字

倉頡輸入法：難人大

注音輸入法：攴ㄨ（音同撲）

輸入完成可顯示：

攴上所施下所效也。☆教效晶𠂇𠂇𠂇
𠂇音勺；𠂇★〇〇人攴攴𠂇子。☆攴𠂇子
見子部。效也。上施故從攴。下效故從
攴𠂇子。古孝切。二部。★凡教之屬皆
〇〇人教。

• 輸入小篆重文「𠂇」

說文解字中收錄的小篆，有部分有異體字形或古文字，這些字形都已收錄在漢字構形資料庫中。欲使用這些字形，可透過漢字構形資料庫取得構字式貼到 Word 文件中，將字形改為小篆，再執行構字式轉換即可。

1. 輸入「教」查詢字形

在字形孳乳視窗上方的文字方塊中輸入「教」，輸入完畢後，按一下「Enter」。字形孳乳下面的方框會顯示所有含「教」的字形，用滑鼠在「教」的字形上按一下，選取「教」的字形。（圖 5.12）

選取字的小篆字形會顯示在「字形演變」的視窗裡。（見次頁，圖 5.13）

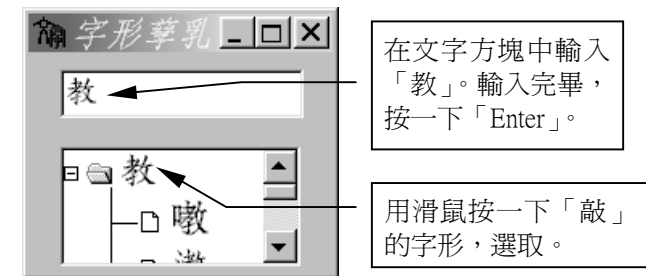


圖 5.12：查詢「教」

1. 複製「𠂔」的構字式

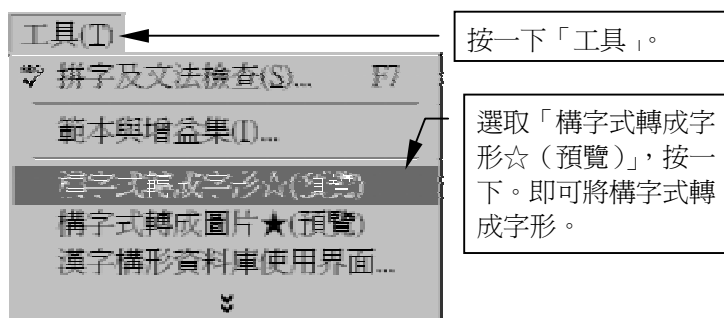
2. 到 Word 裡貼上「𠄎」的構字式

3. 將構字式「    

」的字型改成「北師大小篆」，即可顯示：

注意：若未將欲顯示成小篆的構字式的字型改為「北師大小篆」，執行構字式轉換後，不會顯示小篆字形，只會顯示成楷化字。如「𠂔」，會顯示成「𠂔」。

𡥉 上所施下所效也。☆教效疊韻★从支
𡥉。☆𡥉見子部。效也。上施故從攴。
下效故從𡥉。古孝切。二部。★凡教之
屬皆从教。𡥉 古文教。☆右从古文言★
𡥉 亦古文教。☆從支從攴。★



四、構字式自動校正

構字式轉換完成切換回到「教.doc」的視窗。使用者會發現文件中的構字式和原先輸入的不同（詳見次頁）。

- 原先輸入「教.doc」的文件內容：

- 執行過構字式轉換後「教.doc」的文件內容：

附錄一 缺字的表達方式

使用電腦來處理文件時，會有一些字在電腦裡沒有字形，以致無法輸入，這些字就是缺字。本附錄將說明缺字在電腦中應如何表達。

文件中的缺字要以構字式表達，構字式是由部件和構字符號組成，詳述如下。

一、部件

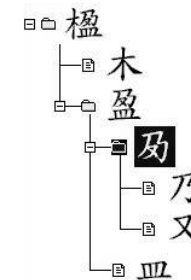
漢字的構形單位稱為部件，部件是有層次的，如：

「木」和「盈」為「楹」的部件

「𠂔」和「皿」為「盈」的部件

「乃」和「又」為「𠂔」的部件

部件的層次如下附錄圖 1.1 所示



附錄圖 1.1 「楹」的構形部件

二、構字符號

構字符號可表達部件連接的方式，也是系統辨識一般文字和缺字的依據。

構字符號有三類，共計十三個，說明及範例如下表：

附錄表 1-1 構字符號表

類別	符號	說明	構字式範例
連接	△	當部件的連接順序由左至右	順＝川△頁
	△	當部件的連接順序由上至下	含＝今△口
	△	當部件的連接順序由外至內	圍＝口△韋
部件序	☐	按部件書寫順序輸入，前後以起始符號（☐）和終止符號（☐）包夾	牖＝ ☐片戶甫☐
	☐		
方便（符號）	∞	二個相同部件直連	炎＝ ∞ 火
	∞	三個相同部件直連	
	∞	二個相同部件橫連	朋＝ ∞ 月
	∞	三個相同部件橫連	
	∞	三個相同部件呈三角狀排列	焱＝ ∞ 火
	∞	四個相同部件橫連	
	∞	四個相同部件直連	
	∞	四個相同部件呈四角狀排列	燚＝ ∞ 火

三、構字式

構字式可由部件連接方式區分成以下三種：

1. 橫連

一個字形中的部件是用橫向的方式連接，字例如：

桃＝木△兆、姚＝女△兆、粥＝弓△米△弓

2. 直連

一個字形中的部件是用直向的方式連接，字例如：

李＝木△子、晁＝日△兆、崋＝山△大△土



3. 包含

一個字形中的部件被另一個部件包含，字例如：



困＝口△木、趟＝走△兆、間＝門△日


爲方便使用者輸入，除了以上三種構字式，還有以下兩種構字式可以使用。

4. 部件序

採用部件連接的方式輸入構字式時，如果遇到不會輸入的部件，可以將該部拆分成更小的部件來輸入。此時只需按照部件的書寫順序輸入，前後以「」和「」包夾即可。







附錄例 1.1：以部件序輸入「綠蠨龜」的「蠨」

輸入「綠虫崔冏龜」。





上例中「蠨」字的構字式爲「虫巒」。但使用者不一定會輸入「巒」，只會輸入「崔」和「冏」。

由於某些字形的部件書寫順序並不明確，爲了兼顧使用上的方便，部件的順序不影響缺字的辨識。

附錄例 1.2：以部件序輸入「囧」

輸入「口井口」、或輸入「口口井」、或輸入「口口井」均可。

5. 方便符號

部分字形會重覆使用到某一部件，如「𣶒」的構字式爲「水水」，可用方便符號「水」來表示，較爲精確。方便符號的表達要列在部件的前面，如「水」=𣶒，不能以「水」來表示。

方便符號可以使用在前述四種構字式中。

附錄表 1-2 方便符號和其他構字式複合的字例

類別	例字	構字式
橫 連	俎	𠂇人  且
直 連	蠶	 无  日   虫
包 含	孱	尸  𠂇子
部件序	𣶒	 𠂇𠂇又酉欠 

附錄二 構字符號的輸入

類別	符號	注音	倉 頡	說 明
連接	⌌	ㄣ	重月(ZB)	形狀像倒下來的B
	⌍		重日(ZA)	形狀像A
	⌎		重金(ZC)	C:contain(包含)
部件序	ㄩ	ㄣ	重竹(ZH)	「形」的倉頡尾碼為「竹」
	ㄩ		重戈(ZI)	點的倉頡尾碼為「戈」
方便 (符號)	∞	ㄣ	竹竹(HH)	H:horizontal (水平的)
	∞		竹竹竹(HHH)	
	∞		竹竹竹竹(HHHH)	
	⊗		難難(XX)	很難輸入
	⊗		難難難(XXX)	
	⊗		難難難(XXX)	
	⊗		難難難難(XXXX)	
	⊗		難難難難(XXXX)	

附錄三 常用部件外字的輸入

部件的倉頡碼，取碼方式為：「難+首尾碼」。「難」表「很難輸入」，也可和一般文字做一區隔。如「𠂇」的倉頡碼為「難口一」。若只有一碼可取，則只要取一碼即可，如「イ」的倉頡碼為「難人」。異寫部件如「イ」為「人」的異寫，用法詳見第三章第四節。

序號	字形	注 音	倉 頡	異寫
1.	丨	ㄍㄨㄣˊ	難中	
2.	丶	ㄗㄨㄣˊ	難戈	
3.	ノ	ㄋㄨㄣˊ	難竹	
4.	ㄥ	ㄣㄣˊ	難山	
5.	ㄣ	ㄣㄣˊ	難弓	
6.	ㄣ	ㄣㄣˊ	難卜	
7.	人	ㄣㄣˊ	難人	人
8.	イ	ㄣㄣˊ	難人	人
9.	冂	ㄣㄣˊ	難月	
10.	冂	ㄣㄣˊ	難月	
11.	ㄣ	ㄣㄣˊ	難戈一	
12.	ㄣ	ㄣㄣˊ	難中弓	刀
13.	ㄣ	ㄣㄣˊ	難心	
14.	ㄣ	ㄣㄣˊ	難尸	
15.	ㄣ	ㄣㄣˊ	難卜	卜
16.	冂	ㄣㄣˊ	難尸中	

序號	字形	注 音	倉 頡	異寫
17.	巳	㇏一セノ	難尸山	
18.	厶	厶	難戈	
19.	夂	㇏ノ	難竹水	
20.	夂	厶メヘ	難竹水	
21.	𠂇	冂一弓ノ	難十	
22.	𠂇	彳セヘ	難山竹	中
23.	𠂇	彳メ弓	難女女	川
24.	𠂇	一么	難女戈	
25.	广	一弓ノ	難戈	
26.	𠂇	一弓ノ	難弓大	
27.	𠂇	㇏一ヘ	難尸一	
28.	𠂇	㇏一ヘ	難弓一	
29.	𠂇	㇏一ヘ	難女一	
30.	𠂇	尸弓	難竹竹	
31.	小	㇏一弓	難心	心
32.	小	㇏一弓	難心	心
33.	扌	尸又ノ	難手	手
34.	支	夕メ	難卜水	
35.	夂	夕メ	難人大	
36.	无	メノ	難一山	
37.	水	尸メヘノ	難水	水
38.	𠂇	尸メヘノ	難水	水
39.	𠂇	尸メセノ	難火	火
40.	𠂇	㇏メ㇏ノ	難月	爪

序號	字形	注 音	倉 頡	異寫
41.	牛	㇏一又ノ	難竹手	牛
42.	𠂇	く口弓ノ	難大竹	犬
43.	正	夕一ノ	難弓一	疋
44.	疒	彳メ尤ノ	難大	
45.	𠂇	夕セ	難弓人	
46.	𠂇	冂メヘ	難田中	目
		メ尤ノ		网
47.	𠂇	尸ヘ	難戈火	示
48.	𠂇	㇏メノ	難竹	竹
49.	糸	冂一ヘ	難女火	糸
50.	𠂇	一尤ノ	難廿土	羊
51.	月	冂又ヘ	難月	月
52.	𠂇	㇏口ノ	難竹難	
53.	𠂇	㇏么ノ	難廿	艸
54.	𠂇	一	難中	衣
55.	𠂇	一㇏ 4	難一田	𠂇
56.	𠂇	尸メノ	難口一	足
57.	𠂇	彳メセヘ	難竹人	
58.	𠂇	彳メセヘ	難卜	
59.	𠂇	一ヘ	難弓中	邑
		冂メヘ		阜
60.	𠂇	彳尤ノ	難尸戈	長
61.	𠂇	夕𠂇 4	難中水	
62.	食	尸ノ	難人戈	食

附錄四 安裝部件外字及輸入法

使用漢字構形資料庫，必需安裝部件外字。部件外字及輸入法的檔案通常會放在 windows 95 或 98 的系統目錄（如 c:\windows）中。

說明：

漢字構形資料庫的部件外字及輸入法的檔案有下列三個

1. 外字檔：cdpeudc.tte
2. 倉頡輸入法：cdpeudcj.tbl
3. 注音輸入法：cdpeudpn.tbl
4. 微軟新倉頡輸入法：ccdpc.tab
5. 微軟新注音輸入法：tcdpc.tab

在安裝漢字構形資料庫時，部件外字及輸入法會同時安裝。

重新安裝部件外字及輸入法

如果之後又安裝了其他的外字檔，導致文件裡的外字對應錯誤。要正確讀取部件外字，就必須重新安裝。

安裝的程序如下：

1. 按一下「開始」，再按一下「程式集」，再按一下「漢字構形資料庫」，選「安裝部件外字及輸入法」，然後會出現以下畫面：



附錄圖 4.1 安裝部件外字及輸入法

2. 直接按「確定」
3. 畫面顯示

安裝完成！

即表示安裝成功，按「確定」結束安裝程序。

4. 重新開機

恢復使用者原先的外字

欲恢復原有的外字及輸入法登錄，登錄的程序如下：

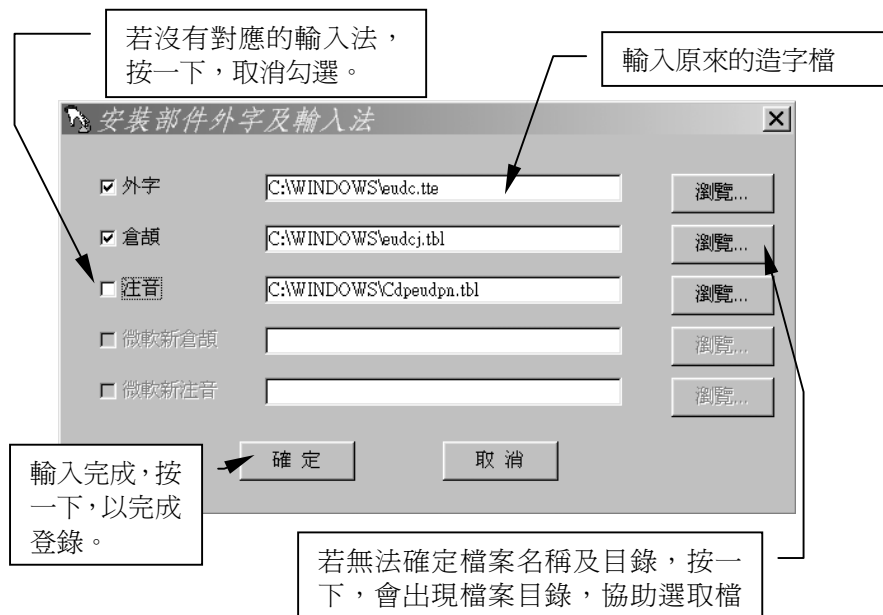
1. 按一下「開始」，再按一下「程式集」，再按一下「漢字構形資料庫」，選「安裝部件外字及輸入法」，然後會出安裝部件外字及輸入法的畫面（附錄圖 4.1）
2. 輸入原先的外字檔（見附錄圖 4.2）。

假設使用者原先的外字檔為 eudc.tte

倉頡輸入法為 edudcj.tbl

無注音輸入法

即可恢復使用者原有的外字檔登錄。



附錄圖 4.2 恢復原有的外字及輸入法