

馬王堆漢墓帛書(貳)-《周易》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
2009/11/10 丁致伶

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 2421 的「傷」字，Unicode 編碼是 3465，由於 XP 並不支援，仍需使用構字式「亻 𠂔易」。

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：

馬王堆漢墓帛書(貳)-《周易》使用舊版造字 55 個，字頻 545 次，這 55 個造字中，14 個可轉成 Windows XP 能顯示的字，字頻 184 次；另外 41 個字必須轉成構字式，字頻 361 次。




轉碼完成製作轉碼分析表，請參考附件一馬王堆漢墓帛書(貳)-《周易》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：檔案內的舊版造字與 Unicode 或漢字構形資料庫所收之字，僅僅只有部件位置或異寫的差異，並無改變該字字義，以 Unicode 字形或漢字構形資料庫所收字形的構字式取代。例：造字編號 5230 的「土畝且」字，是 Unicode 字形「壇」的異體字，僅僅只有部件位置或異寫的差異，並無改變該字字義，因此以 Unicode 字形「壇」取代。
2. 待造字：Unicode 及漢字構形資料庫皆未收錄的舊版造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 1301 的「」。
3. 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 4158 的「譎」字，Unicode 字型呈現為「譎」，實際上仍為同一字。

附件一、馬王堆漢墓帛書(貳)-《周易》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
296		82	FBEC				
297		47	FBED				
298		9	FBEE				
299		13	FBEF				
300		13	FBF0				

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
308	∞	1	FBF8			∞	
379	支	1	FCA2			支	
393	尿	2	FCB0			尿	
404	弟	1	FCBB			弟	
528	長	4	FD78			長	
535	亾	1	FDA1			亾	
570	寔	7	FDC4			寔	
595	卓	4	FDDD			卓	
613	繸	2	FDEF			繸	
629	巳	1	FE40			巳	
637	戔	2	FE48			戔	
679	叢	1	FE72			叢	
684	一	1	FE77			一	
709	弟	2	FEB2			弟	
757	乚	1	FEE2			乚	
761	夂	1	FEE6			夂	
766	糸	3	FEEB			糸	
768	艹	8	FEED			艹	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
770	扌	5	FEEF			扌	
775	ㄥ	1	FEF4			ㄥ	
780	讠	4	FEF9			讠	
1057	夔	1	8FD4	3687		夔	
1238	戾	5	90EC	77E6	戾		
1301		1	916C			亻 厶 匈	待造字。
1886	榦	1	9541			車 厶 余	
2181	訖	2	96ED	8A26	訖		
2421	傷	1	98A3	3465		亻 厶 易	
2618	缶	2	99CB	26222		缶	缶，異體字問題。
2836	三	16	9B49			三	
2837	三	17	9B4A			三	
3008		1	9C58			土 厶 显	埏，異體字問題。
3028	礻	1	9C6C			礻	
3068	三	18	9CB6			三	
3069	三	15	9CB7			三	
3071	三	18	9CB9			三	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3072	三	16	9CBA			三	
3073	三	16	9CBB			三	
3074	三	16	9CBC			三	
3840	坂	1	81A9	5742	坂		
4100	菘	3	8351	82FD	菘		
4158	譌	1	83AD	8B4C	譌		Unicode 字型 呈現差異。
4286	从	1	846E	4ECE	从		
4496	段	4	85C5	53DA	段		
5175	耜	1	89F8	800A	耜		
5230		1	8A70			形土畝 且	壇，異體字問 題。
5429	拇	2	8BBC	6802	拇		
5516		1	8C54			∞束	棘，異體字問 題。
5534	塢	6	8C66	4CA8		工 鳥	
5619	繚	1	8CDD	470C		繚	
5860	无	159	C6D3	65E0	无		

二、手動取代表說明：

馬王堆漢墓帛書(貳)-《周易》檔案中需手動修改的情況分為以下幾種：

(一) 馬王堆漢墓帛書(貳)-《周易》檔案中出現「★」字形，共計 4 字，與語言所討論後以 Unicond 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. 頁數：「★」字形所在位置之原書頁數。
2. 篇名：「★」字形所在位置之原書篇名。
3. 檔案原文摘錄：摘錄馬王堆漢墓帛書(貳)-《周易》檔案有「★」字形的文句。
4. 修改字形：與語言所討論後決定的「★」之字形，以此字取代檔案原文的「★」。
5. 備註：記錄修改相同詞句的次數或其他事項。

附件二、《馬王堆漢墓帛書(貳)-周易》「★」字手動取代表

頁數	篇名	檔案原文摘錄	修改字形	備註
p. 52	訟第五	終朝三★(禡)	𠄎扌肅一□	待造字
p. 111	損第十二	★(曷)之用	𠄎久口內□	待造字
p. 284	解第三十	公用射★(隼)	𠄎隼	待造字
p. 385	夬第四十二	覓★(陸)缺(夬)	刺隼	待造字

(二) 馬王堆漢墓帛書(貳)-《周易》檔案中，對於舊版造字以外發現的錯字，以手動修改，並製作錯字修改表，請參考附件三。欄位說明如下：

1. 檔案原文摘錄：摘錄檔案馬王堆漢墓帛書(貳)-《周易》中的文句段落，其中紅字底線的部分，為需要手動取代的原文。
2. 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。

3. 備註：記錄修改相同詞句的次數或其他事項。

附件三、馬王堆漢墓帛書(貳)-《周易》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
