

《入唐求法巡禮行記》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
2008/4/30 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 2561 的「𠄎」字，Unicode 編碼是 279AA，由於 XP 並不支援，仍需使用構字式「功𠄎言」。

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：

《入唐求法巡禮行記》使用舊版造字 74 個，字頻 701 次，這 74 個造字中，68 個可轉成 Windows XP 能顯示的字，字頻 693 次；另外 6 個字必須轉成構字式，字頻 8 次。

轉碼完成製作轉碼分析表，請參考附件一《入唐求法巡禮行記》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**：為了使用者查詢和使用的方便，在處理異體字時最主要的方式是以標準字取代，除非是專有名詞或特殊情形，如：人名、地名等。例：造字編號 1513 的「粮」字，是「糧」的異體字，以標準字「糧」取代。
2. **錯字，以程式全部取代**：與原書字形不符，並查詢教育部異體字字典確認非異體字後，皆歸類為錯字，以程式取代為正確字形。例：檔案中使用造字編號 2461 的「葱」字，而原書使用之字為「葱」，查詢教育部異體字字典確認「葱」並非「葱」之異體字，因此歸為錯字，以程式取代為正確字形「葱」。
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 4091 的「臯」字，Unicode 字型呈現為「臯」，實際上仍為同一字。

附件一、《入唐求法巡禮行記》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
842	𪗇	1	8E78	3AB7		𪗇	
887	鷲	1	8EC7	9E0E	鷲		鷲，異體字問題。
1513	粮	6	92C5	7CAE	粮		糧，異體字問題。
1530	糺	2	92D6	7CFA	糺		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1784	萇	4	9478	855A	萇		
1978	讚	3	95BF	8B83	讚		
2342	穉	1	97F1	4176	禾厶辱		
2461	葱	4	98CB	44E4	艹厶忽		葱，錯字，以程式全部取代。
2524	饒	1	994B	994D	饒		
2561	訃	1	9970	279AA	功厶言		
2828	蝸	2	9B41	8744	蝸		
3805	勅	201	8164	52C5	勅		
3807	驗	33	8166	9A10	驗		
3814	却	85	816D	5374	却		
3822	叁	2	8175	53C1	叁		
3824	叙	2	8177	53D9	叙		
3840	坂	27	81A9	5742	坂		
3843	堦	2	81AC	5826	堦		
3860	尠	2	81BD	5C20	尠		
3864	峯	13	81C1	5CEF	峯		峰，異體字問題。
3883	徧	1	81D4	5FA7	徧		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3891	慙	1	81DC	6159	慙		
3893		1	81DE		亻 ㄩ 妄		待造字
3923		1	81FC		晶 ㄩ 宜		疊，異體字問題。
3931	椀	5	8245	6900	椀		
3940	鬱	1	824E	6B1D	鬱		
3966	烟	2	8268	70DF	烟		
3978	猪	1	8274	732A	猪		豬，異體字問題。
3984	瑠	10	827A	7460	瑠		
3996	踈	3	82A8	758E	踈		
4016	着	89	82BC	7740	着		
4038	竈	2	82D2	7AC8	竈		
4041	豎	1	82D5	7AEA	豎		
4080	羣	4	82FC	7FA3	羣		
4089	脇	3	8346	8107	脇		
4090	脚	18	8347	811A	脚		腳，異體字問題。
4091	臯	5	8348	81EF	臯		Unicode 字型呈現差異。
4100	苳	1	8351	82FD	苳		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4105	菓	8	8356	83D3	菓		
4137	袜	1	8376	889C	袜		
4146	覩	1	83A1	89A9	覩		睹，異體字問題。
4167	賚	1	83B6	8CEB	賚		
4176	輒	2	83BF	8F19	輒		
4184	迹	9	83C7	8FF9	迹		
4193	鈎	2	83D0	920E	鈎		
4195	鉢	11	83D2	9262	鉢		
4203	隣	2	83DA	96A3	隣		
4218	菲	5	83E9	97EE	菲		
4226	凜	7	83F1	98E1	凜		
4240	髯	2	8440	9AF4	髯		
4256	鷄	1	8450	9DC4	鷄		雞，異體字問題。
4307	罵	2	84A5	99E1	罵		罵，異體字問題。
4460	勅	1	85A1	52D1	勅		
4515	呪	1	85D8	546A	呪		
4519	咲	2	85DC	54B2	咲		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4622	該	1	86A6	8A43	該		
4738	睺	1	875B	777A	睺		
4856	疊	1	87F3	3CB2	疊厶毛		
4904	尅	4	8864	5C05	尅		
5099	惣	58	89AC	60E3	惣		
5101	馱	3	89AE	4B7E	馬厶犬		
5165		1	89EE		晶厶宜		疊，異體字問題。
5176		2	89F9		冫厶冗		沉，異體字問題。
5179	駟	1	89FC	99C8	駟		
5300	鬪	3	8AD8	9B2A	鬪		
5434	湟	1	8BC1	6E7C	湟		
5542	鑠	1	8C6E	93C1	鑠		
5551	賈	1	8C77	8CF7	賈		
5558	龕	13	8C7E	9E84	龕		
5583	殮	2	8CB9	98F1	殮		
5678	量	2	8D59	757A	量		
5704	氈	3	8D73	6C0A	氈		氈，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5742	澁	1	8DBB	6F81	澁		
5753	潛	2	8DC6	6F5C	潛		

二、手動取代表說明：

《入唐求法巡禮行記》檔案中需手動修改的情況分為以下幾種：

(一) 《入唐求法巡禮行記》檔案中出現「●」字形，共計 57 字，比對原書後以 Unicod 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. 頁碼：「●」字形所在位置之原書頁碼。
2. 檔案原文摘錄：摘錄《入唐求法巡禮行記》檔案有「●」字形的文句。
3. 原書字形：「●」在原書中之字形，以此字取代檔案原文的「●」。
4. 備註：記錄修改相同詞句的次數或其他事項。

附件二、《入唐求法巡禮行記》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 2	沙埋●	樛	
p. 2	●。眾人設謀	木厶我	待造字
p. 2	結纜●	樛	
p. 2	●。亥時	木厶我	待造字
p. 5	大使等為●漂勞	憇	
p. 9	水手丁勝小●	磨	
p. 11	楊隻糸●等趨來	磨	
p. 15	許●皇太子	𠂇厶文	

頁碼	檔案原文摘錄	原書字形	備註
p. 15	皇太子擬●父王	𠂇 𠂇 𠂇	
p. 15	仍父王●己子	𠂇 𠂇 𠂇	
p. 17	計●宰相及大官	𠂇 𠂇 𠂇	
p. 17	亂●計萬人已上	𠂇 𠂇 𠂇	
P. 40	同心●害	𠂇 𠂇 𠂇	
p. 41	鉢盂、●	食 𠂇 周	待造字
p. 41	●、文書	食 𠂇 宛	
p. 45	當大●正中	擢	
p. 45	舳小●腋門	擢	
p. 57	亦官私絹，●顯	糸 𠂇 頰	待造字
p. 61	暫請●艇	舟 𠂇 遊	待造字
p. 63	●艇一雙	舟 𠂇 遊	有 2 處 待造字
p. 63	補綴●艇	舟 𠂇 遊	待造字
p. 78	●奪不少	𠂇 𠂇 𠂇	
p. 81	伏蒙●前大守	筑	
p. 83	城中●却四千餘人	𠂇 𠂇 𠂇	
p. 86	邀上●裏啜茶	形 厂 耳 惠 口	改為廳
p. 123	獄卒現前忿●	忄 𠂇 勃	待造字
p. 131	被人藥●	𠂇 𠂇 𠂇	
p. 134	自出帔●子一領	糸 𠂇 奧	
p. 148	●頭足	食 𠂇 念	
p. 155	准勅盡斬●訖	𠂇 𠂇 𠂇	
p. 160	公主在路自●之	𠂇 𠂇 𠂇	
p. 160	令●天下摩尼師	𠂇 𠂇 𠂇	
p. 160	作沙門形而●之	𠂇 𠂇 𠂇	

頁碼	檔案原文摘錄	原書字形	備註
p. 172	盡●破家	𠂇𠂇文	
p. 174	拙●除淮南節度使	土𠂇宗	待造字
p. 175	斬●破家	𠂇𠂇文	
p. 175	打●三百餘	𠂇𠂇文	
p. 176	並當處打●	𠂇𠂇文	
p. 176	三長月不許●命	𠂇𠂇文	
p. 177	兩軍兵馬圍着●之	𠂇𠂇文	
p. 179	令進藥酒而藥●矣	𠂇𠂇文	
p. 179	天子索弓射●	𠂇𠂇文	
p. 179	不合●之	𠂇𠂇文	
p. 179	斬●六七度也	𠂇𠂇文	
p. 179	當時打●	𠂇𠂇文	
p. 180	無故射●虞候	𠂇𠂇文	
p. 180	仰諸寺營辦床席●	毛△曾	
p. 181	百司●●不已	忄△念忄△念	
p. 184	當時決●者	𠂇𠂇文	
p. 195	當處決●	𠂇𠂇文	
p. 200	到萊州界●山	山△牢	
p. 200	今且在●	山△牢	
p. 201	便船往●山	山△牢	
p. 201	其後却擬向●山	山△牢	
p. 201	到●山南椒家莊	山△牢	