

《出曜經》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
2009/04/22 丁玟伶

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 4856 的「𪛗」字，Unicode 編碼是 3CB2，由於 XP 並不支援，仍需使用構字式「疊𪛗毛」。

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：

《出曜經》使用舊版造字 106 個，字頻 826 次，這 106 個造字中，95 個可轉成 Windows XP 能顯示的字，字頻 793 次；另外 11 個字必須轉成構字式，字頻 33 次。

轉碼完成製作轉碼分析表，請參考附件一《出曜經》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 4134 的「衛」字，是「衛」的異體字，以標準字「衛」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 4561 的「𠄎」字，而原書使用之字為「鹿」，查詢教育部異體字字典確認「𠄎」並非「鹿」之異體字，因此歸為錯字，以程式取代為正確字形「鹿」。
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 3978 的「猪」字，Unicode 字型呈現為「猪」，實際上仍為同一字。

附件一、《出曜經》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
845	𧈧	1	8E7B	9F26	𧈧		疊 𠄎 毛，錯字，以程式全部取代。
888	捺	6	8EC8	3B88		木 𠄎 奈	
1024	虵	1	8FB3	8675	虵		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1032	憊	8	8FBB	2039D		𠂇𠂇 心	
1091	漚	8	8FF6	58CD	漚		
1237	稍	1	90EB	77DF	稍		
1400	窓	1	91F1	7A93	窓		
1536	麤	3	92DC	9E81	麤		
1978	讚	5	95BF	8B83	讚		讚，異體字問題。
1981	讐	3	95C2	8B90	讐		讐，異體字問題。
2044	踈	2	9642	8E08	踈		
2190	稔	1	96F6	6BFA	稔		
2278	銓	4	97B1	936E	銓		
2505	翻	2	98F7	98DC	翻		
2516	餽	1	9943	9919	餽		
2631	鴉	4	99D8	9D44	鴉		
2885	叫	1	9B7A	544C	叫		
2969	諧	3	9BF0	8B5B	諧		
3143	潰	1	9D42	6FFD	潰		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3801	剗	1	8160	5257	剗		
3804	効	3	8163	52B9	効		
3805	勅	29	8164	52C5	勅		
3808	勾	1	8167	5303	勾		
3814	却	11	816D	5374	却		
3817	廁	2	8170	53A0	廁		廁，異體字問題。
3850	姪	1	81B3	59D9	姪		
3860	尠	14	81BD	5C20	尠		
3876	廻	1	81CD	5EFB	廻		
3891	慙	4	81DC	6159	慙		
3897	慙	1	81E2	615C	慙		
3920	皓	2	81F9	6667	皓		
3940	藟	11	824E	6B1D	藟		
3955	澗	6	825D	3D4E		汜 㒰 澗	澗，異體字問題。
3966	烟	3	8268	70DF	烟		
3976	犁	1	8272	7282	犁		
3978	猪	14	8274	732A	猪		Unicode 字型

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							呈現差異。
3980	獾	10	8276	7328	獾		
3984	瑠	2	827A	7460	瑠		
3992	嘗	2	82A4	751E	嘗		
3996	疎	1	82A8	758E	疎		
4001	瘥	4	82AD	7602	瘥		
4012	鼓	2	82B8	76B7	鼓		鼓，異體字問題。
4028	稟	4	82C8	7980	稟		
4031	杭	1	82CB	79D4	杭		
4035	穽	1	82CF	7A7D	穽		
4038	竈	1	82D2	7AC8	竈		
4041	豎	13	82D5	7AEA	豎		
4045	箒	2	82D9	7B92	箒		
4068	綫	3	82F0	7DAB	綫		
4085	耻	1	8342	803B	耻		
4089	脇	8	8346	8107	脇		
4090	脚	10	8347	811A	脚		
4105	菓	32	8356	83D3	菓		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4107	葱	1	8358	8471	葱		
4134	衛	100	8373	885E	衛		衛，異體字問題。
4140	衽	1	8379	88B5	衽		
4146	覩	101	83A1	89A9	覩		睹，異體字問題。
4173	駝	5	83BC	8EAD	駝		
4176	輒	2	83BF	8F19	輒		
4184	迹	30	83C7	8FF9	迹		
4193	鈎	2	83D0	920E	鈎		鈎，異體字問題。
4195	鉢	60	83D2	9262	鉢		
4203	隣	11	83DA	96A3	隣		
4226	飡	1	83F1	98E1	飡		
4237	髡	1	83FC	9AE0	髡		
4256	鷄	3	8450	9DC4	鷄		雞，異體字問題。
4267	鼈	6	845B	9F08	鼈		
4342	倮	10	84C8	502E	倮		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4398	贛	4	8541	4D43		麥厶黃	
4416	汧	2	8553	44D1		灬厶汧	
4448		8	8573			形糸厂 墨口	纏，異體字問題。
4495	尪	4	85C4	9EA8	尪		Unicode 字型呈現差異。
4496	段	3	85C5	53DA	段		
4515	呪	39	85D8	546A	呪		
4558	廐	1	8644	53A9	廐		廐，異體字問題。
4561		1	8647			形乚厂 曲比口	廐，錯字，以程式全部取代。
4566	礪	1	864C	78F5	礪		
4693	隴	2	86ED	22931		亻厶龍	
4696	燿	1	86F0	3E0C		火厶霍	
4723	塶	3	874C	58C3	塶		
4739	鞞	1	875C	97DB	鞞		
4768	妬	17	8779	59AC	妬		妒，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4775	妣	22	87A2	59B7	妣		
4856	毳	4	87F3	3CB2		疊厶毛	
4889	毳	1	8855	3CAA		登厶毛	
4904	尅	3	8864	5C05	尅		
5002	鞞	2	88E8	4A6D		鞞	
5132	愜	1	89CD	6198	愜		
5133	愜	1	89CE	6199	愜		
5169	危	5	89F2	6239	危		
5188	迺	1	8A46	62EA	迺		
5290	携	1	8ACE	643A	携		
5300	鬪	68	8AD8	9B2A	鬪		
5347	憾	1	8B48	617D	憾		
5364	殂	1	8B59	239D1		歹厶凶	
5478	飴	1	8BED	4D77	飴		
5542	鑠	2	8C6E	93C1	鑠		
5559	欸	4	8CA1	6B35	欸		
5587	欸	2	8CBD	6B58	欸		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5588	恠	2	8CBE	6061	恠		
5697	煑	1	8D6C	7151	煑		
5702	厨	15	8D71	53A8	厨		
5742	澁	2	8DBB	6F81	澁		
5753	潜	5	8DC6	6F5C	潜		
5799	蜀	1	8DF4	262FA		𠄎𠄎𠄎	
5802	慄	1	8DF7	5E56	慄		

二、手動取代表說明：

《出曜經》檔案中出現「●」字形，共計106字，比對原書後以Unicode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《出曜經》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、《出曜經》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 609	●悔心	悋	
p. 610	無所●惜	悋	

頁碼	檔案原文摘錄	原書字形	備註
p. 612	●脹臭爛	臙	
p. 612	鴿蛆蟲●	口𠂔𠂔	
p. 612	●臭穢難近	口𠂔束	
p. 613	人●舉號	𠂔	有 2 處
p. 613	已生在外●	歹𠂔冬	
p. 616	無失●體	身𠂔丘	
p. 617	飲食●樂	食𠂔燕	待造字
p. 617	飲食歡*●	食𠂔燕	待造字
p. 617	車●馬璫	形王巨木	有 11 處
p. 620	●·額如油光	醫	
p. 620	新●牛乳	殼𠂔牛	待造字
p. 621	飛鳥●河白	壽𠂔鳥	
p. 622	指刺不●	而𠂔火	
p. 622	●詣塚墳	𠂔	
p. 624	大水暴●駛流	漲	
p. 625	為蠅所嚙●	口𠂔束	
p. 634	●草生生	形** 呂女干	
p. 635	施●柯柯長	金𠂔斧	
p. 635	●往斫樹	金𠂔斧	
p. 636	是所謂●生	疤	
p. 636	終生●蟲	疤	

頁碼	檔案原文摘錄	原書字形	備註
p. 639	禪●法杖	𠄎	
p. 639	以●往擊得	𠄎	
p. 644	二女●五百	𠄎△牛	待造字
p. 644	復●二百	𠄎△牛	待造字
p. 644	復●百二十五	𠄎△牛	待造字
p. 644	●六十牛飲	𠄎△牛	待造字
p. 644	復●三十飲	𠄎△牛	待造字
p. 644	●十五飲	𠄎△牛	待造字
p. 644	七復●七·	𠄎△牛	待造字
p. 645	諸●子安隱居	𠄎△卵	
p. 645	受不●其義	恪	
p. 645	真珠車●	𠄎王巨木□	
p. 647	水性龜●之	魚△龜	
p. 655	悔●·是故	恪	
p. 657	不曾麤●	麥△廣	
p. 663	吐言麤●	麥△廣	
p. 664	●大如芥子	庖	
p. 665	生●瘡即取	庖	
p. 665	因姪嫉致●	恪	
p. 667	饑形●極苦	繼△足	待造字
p. 668	則受其●	𠄎イ ∞无心□	

頁碼	檔案原文摘錄	原書字形	備註
p. 669	身體●脹	臃	
p. 671	如●牛乳	殼△牛	待造字
p. 671	如●牛*乳	殼△牛	待造字 有 2 處
p. 676	致不●濫人	枉	
p. 676	者●濫人物	枉	
p. 677	●無復生	𠄎△朋	待造字
p. 678	譬如執●草	𠄎 ^𠄎 𠄎 女 干 □	
p. 678	*●草	𠄎 ^𠄎 𠄎 女 干 □	有 4 處
p. 685	偽●殺生類	枉	
p. 690	●豆食之	豆△𦉳	
p. 690	苦酒●豆食	豆△𦉳	
p. 690	●*桃甘蔗	𠄎 ^𠄎 木 甫 □	待造字
p. 691	自不恃德●	菱△欠	
p. 693	心無慳●	恪	
p. 700	●芽得生	𠄎 ^𠄎 △ 朋	待造字 有 2 處
p. 703	豈不●乎	枉	
p. 703	坐地●哭	口△號	待造字
p. 707	治牢固●	掌△支	待造字
p. 709	以鼻●之	𠄎 ^𠄎 口 自 死 □	待造字

頁碼	檔案原文摘錄	原書字形	備註
p. 710	是剛●石裏	鞭	
p. 715	不懷慳●	恪	
p. 719	●哭喚呼	口𠂔罩	
p. 726	無所●惜也	恪	
p. 732	何為●叫言	口𠂔罩	
p. 735	六情●塞不通	閉	
p. 742	●蔗葡	甘	
p. 746	內心●惜	恪	
p. 747	●要品第三十	兩𠂔隻	有 2 處
p. 747	夜光便黤●	黑𠂔炎	
p. 747	顏色黤*●	黑𠂔炎	
p. 747	便黤*●也	黑𠂔炎	有 2 處
p. 749	生●蔗之	甘	
p. 749	多食瞪●	目𠂔夢	
p. 752	●有形之類	蚤	
p. 753	非法濫●百姓	枉	
p. 756	●八輩	兩𠂔隻	
p. 756	不●物理	枉	
p. 756	不●人物	枉	
p. 756	具毳毳●	翕𠂔毛	
p. 761	●哭受其報	口𠂔罩	

頁碼	檔案原文摘錄	原書字形	備註
p. 761	鐵丸●哭	口𠂔罩	
p. 761	●哭受報	口𠂔罩	有 2 處
p. 772	●飛蠕動	𠂔虫日月◻	待造字