

《論衡校釋》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
2008/12/08 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 961 的「枅」字，Unicode 編碼是 3B4A，由於 XP 並不支援，仍需使用構字式「木𠂔片」。

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：

《論衡校釋》使用舊版造字 168 個，字頻 1193 次，這 168 個造字中，146 個可轉成 Windows XP 能顯示的字，字頻 1127 次；另外 22 個字必須轉成構字式，字頻 66 次。

轉碼完成製作轉碼分析表，請參考附件一《論衡校釋》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 982 的「胃」字，是「胸」的異體字，以標準字「胸」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 5505 的「耆」字，而原書使用之字為「耆」，查詢教育部異體字字典確認「耆」並非「耆」之異體字，因此歸為錯字，以程式取代為正確字形「耆」。
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 3935 的「槩」字，Unicode 字型呈現為「槩」，實際上仍為同一字。
4. 待造字：Unicode 及漢字構形資料庫皆未收錄的舊漢籍造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 5341 的「⁺⁺𠄎

附件一、《論衡校釋》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
814	墓	2	8E5C	87C7	墓		
826	癱	5	8E68	7655	癱		
827	鮫	2	8E69	9B8C	鮫		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
829	鼻	9	8E6B	81F0	鼻		
832	淹	3	8E6E	6CB2	淹		
850	坑	7	8EA2			土厶元	
899	馳	1	8ED3	99DE	馳		
915	遲	1	8EE3	905F	遲		
931	磕	1	8EF3	791A	磕		
961	枅	5	8F52	3B4A		木厶片	
967	僂	2	8F58	50AB	僂		
972	犇	7	8F5D	7287	犇		
977	磨	2	8F62	9E8F	磨		
978	戴	1	8F63	4CD2		戈厶鳥	
982	胷	5	8F67	80F7	胷		胸，異體字問題。
986	犴	1	8F6B	7302	犴		
1024	虵	94	8FB3	8675	虵		
1053	傲	2	8FD0	6160	傲		
1067	碁	1	8FDE	68CA	碁		
1091	漣	1	8FF6	58CD	漣		
1226	罰	20	90E0	7F78	罰		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1238	疾	1	90EC	77E6	疾		
1354	裸	1	91C3	79A9	裸		
1591	羗	1	9354	7F97	羗		
1642	脉	29	93A9	8109	脉		
1668	臈	1	93C3	81C8	臈		
1864	蛄	1	94EA	872B	蛄		
1885	蠶	1	9540	882D	蠶		
2044	踈	5	9642	8E08	踈		
2053	躡	1	964B	8EA7	躡		
2179	胛	1	96EB	26643		胛	
2323	質	4	97DE	287CA		質 厶 卩	
2439	礪	4	98B5	7920		石 厶 慈	
2501	颯	3	98F3	98C3	颯		
2506	飭	1	98F8	991D	飭		
2564	剌	8	9973	34E8		夾 厶 刂	
2574	勵	1	997D	3899		广 厶 勤	
2651	鸚	14	99EC	9E1C	鸚		
3001		1	9C51	87A5	螿		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3048	叅	2	9CA2	53C5	叅		
3787	冢	7	8152	51A2	冢		
3791	况	1	8156	51B5	况		况，異體字問題。
3794	湊	2	8159	51D1	湊		湊，異體字問題。
3804	効	1	8163	52B9	効		
3807	驗	1	8166	9A10	驗		
3814	却	23	816D	5374	却		
3817	廁	2	8170	53A0	廁		廁，異體字問題。
3819	廝	1	8172	53AE	廝		廝，異體字問題。
3826	叶	3	8179	53F6	叶		
3835	噍	2	81A4	564D	噍		
3840	坂	2	81A9	5742	坂		
3843	堦	1	81AC	5826	堦		
3850	姪	9	81B3	59D9	姪		
3853	褻	4	81B6	2776C		敝衤衣	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3855	孽	1	81B8	5B7C	孽		
3858	冤	15	81BB	5BC3	冤		冤，異體字問題。
3868	卮	1	81C5	5DF5	卮		
3883	徧	5	81D4	5FA7	徧		
3891	慙	10	81DC	6159	慙		
3894	筋	8	81DF	4225		竹 厶 勹	
3921	皙	17	81FA	6673	皙		
3930	捰	2	8244	6901	捰		
3935	槩	1	8249	3BA3	槩		Unicode 字型呈現差異。
3937	檝	7	824B	6A9D	檝		
3951	湏	1	8259	6E4F	湏		
3966	烟	2	8268	70DF	烟		
3967	炳	1	8269	712B	炳		
3974	牀	3	8270	7240	牀		
3978	猪	6	8274	732A	猪		Unicode 字型呈現差異。
3996	疎	2	82A8	758E	疎		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4004	瘠	1	82B0	763B	瘠		
4008	癯	1	82B4	766F	癯		
4010	皐	13	82B6	7690	皐		皐，異體字問題。
4012	鼓	10	82B8	76B7	鼓		
4015	眦	1	82BB	7726	眦		
4016	着	23	82BC	7740	着		
4028	稟	2	82C8	7980	稟		
4034	穉	4	82CE	7A49	穉		
4038	竈	15	82D2	7AC8	竈		
4040	竝	3	82D4	7ADD	竝		
4041	豎	2	82D5	7AEA	豎		
4044	筋	2	82D8	7B6F	筋		
4051	籊	2	82DF	7C35	籊		
4052	簾	1	82E0	7C34	簾		Unicode 字型呈現差異。
4055	粃	1	82E3	7C83	粃		
4063	藜	2	82EB	7CF5	藜		
4077	罇	5	82F9	7F47	罇		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4079	待	1	82FB	38E5		彳厶侍	
4080	羣	38	82FC	7FA3	羣		
4085	耻	6	8342	803B	耻		
4091	臯	8	8348	81EF	臯		
4095	舩	8	834C	8229	舩		
4100	菘	2	8351	82FD	菘		
4107	葱	7	8358	8471	葱		
4109	葛	1	835A	84AD	葛		
4114	藜	1	835F	8616	藜		
4115	藁	1	8360	85C1	藁		
4132		196	8371			血厶丞	眾，異體字問題。
4134	衛	50	8373	885E	衛		衛，異體字問題。
4139	袴	2	8378	88B4	袴		
4140	衽	2	8379	88B5	衽		
4142	裛	27	837B	8943	裛		裛，異體字問題。
4146	覩	4	83A1	89A9	覩		睹，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4158	譌	1	83AD	8B4C	譌		
4167	賚	3	83B6	8CEB	賚		
4174	躑	4	83BD	8EB6	躑		
4176	輒	24	83BF	8F19	輒		
4180	臯	1	83C3	8FA0	臯		
4184	迹	17	83C7	8FF9	迹		
4186	遡	1	83C9	9061	遡		
4190	邳	1	83CD	90C4	邳		
4218	葦	3	83E9	97EE	葦		
4226	凜	7	83F1	98E1	凜		
4237	髡	10	83FC	9AE0	髡		
4242	鬪	17	8442	9B2D	鬪		
4256	鷄	22	8450	9DC4	鷄		雞，異體字問題。
4259	麇	1	8453	9E95	麇		
4260	麇	26	8454	9E9E	麇		
4265	鼃	1	8459	9F03	鼃		
4266	鼃	6	845A	9F02	鼃		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4267	鼈	16	845B	9F08	鼈		
4307	罵	1	84A5	99E1	罵		罵，異體字問題。
4309	刼	1	84A7	523C	刼		
4342	倮	15	84C8	502E	倮		
4371	倜	1	84E5	50F4	倜		
4455	劒	7	857A			僉 ㄩ 刃	
4456	妨	3	857B	8684	妨		
4458	勑	1	857D	52CC	勑		
4460	勅	14	85A1	52D1	勅		
4507	吳	60	85D0	5434	吳		吳，異體字問題。
4533	唸	1	85EA	552B	唸		
4643	徃	1	86BB	5F83	徃		
4700	埕	2	86F4	586A	埕		
4768	妬	10	8779	59AC	妬		妒，異體字問題。
4902	兗	1	8862	5156	兗		
4904	尅	1	8864	5C05	尅		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4996	斲	6	88E2	65B5	斲		
5031	閔	1	8946	95B4	閔		
5050	熿	1	8959	71CC	熿		
5105	埴	1	89B2	5860	埴		
5133	憇	1	89CE	6199	憇		
5147	恠	1	89DC	6060	恠		
5162	嶷	1	89EB	37EA		山厶畏	
5176		11	89F9			氵厶冗	沉，異體字問題。
5244	嫫	1	8A7E	3738		薛厶女	
5248		2	8AA4			扌厶間	待造字。
5253		7	8AA9			卜厶冂 厶冂	高，異體字問題。
5265	藟	2	8AB5	26E1C		艹厶艹 厶回	
5302	昏	3	8ADA	662C	昏		
5310	烏	2	8AE2	8204	烏		
5320	瘡	1	8AEC	24E07		疒厶夊	
5335	蚩	1	8AFB	459F		亡厶虫	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5339	奸	7	8B40	59E7	奸		
5341		1	8B42			𠄎 𠄎 𠄎	待造字。
5439	栢	2	8BC6	6822	栢		
5443	跣	1	8BCA	27FF6		足 𠄎 屯	
5459	缺	1	8BDA	21647		垂 𠄎 夬	
5482	璽	4	8BF1	8812	璽		
5505	耆	1	8C49	264BE		老 𠄎 占	耆，錯字，以程式全部取代。
5542	鑠	1	8C6E	93C1	鑠		
5551	賣	3	8C77	8CF7	賣		
5578	弛	4	8CB4	38AE		弓 𠄎 屯	
5588	恠	1	8CBE	6061	恠		
5697	煑	1	8D6C	7151	煑		
5702	厨	4	8D71	53A8	厨		
5706	敍	3	8D75	654D	敍		敍，異體字問題。
5752	摠	1	8DC5	63D4	摠		
5860	无	1	C6D3	65E0	无		

二、手動取代表說明：

《論衡校釋》檔案中出現「●」字形，共計60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《論衡校釋》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、《論衡校釋》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 11	蠶●之黨	萬△虫	
p. 28	王霸同一●	业△業	改為業
p. 91	又●(俱)誅死	惧	
p. 102	遭●適履	尸△立	
p. 104	賊貨而命妨●也	殷△禾	
p. 104	●也·世謂宅有吉凶	殷△禾	
p. 163	乃為「女」旁「●」	臣	
p. 170	不見●(埵)螺(堞)	虫△垂	待造字
p. 320	猶●岬也	山△突	
p. 323	食合●	利△虫	
P. 397	蓋起問難此(●)說	言△比	
p. 524	上儉，己不●	形△人方△	

頁碼	檔案原文摘錄	原書字形	備註
p. 549	熱究(突)浹(●)	突	
p. 554	將使●	𠂇 𠂇 爽	待造字
p. 554	(●)然各知所	爽	
p. 603	經徒所摘(●)	摺	
p. 649	是故夏末蜻●鳴	列 虫	
p. 689	小而緩者用鈴●	𠂇 𠂇 狄	待造字
p. 689	(●)，彰	簌	
p. 710	●●閩亡 虫 皆食人	虫 𠂇 造 虫 𠂇 虱	
p. 716	●(癥)蝦(瘕)有蟲	虫 𠂇 微	待造字
p. 760	鮭●(●)	角 𠂇 虎 角 𠂇 虜	有 10 處
p. 801	兩揚軒(軒)●之聲	車 𠂇 蓋	
p. 873	雖虎●勇猥	𠂇 𠂇 儿	
p. 888	太子●(使)登僕車	趋	
p. 898	(●)水擊其墓	欒	
p. 900	故使●水見之	欒	
p. 955	鮭與●	魚 𠂇 多	
p. 955	●，故人食鮭肝	魚 𠂇 叔	
p. 955	為●	魚 𠂇 多	
p. 955	●螫有毒	魚 𠂇 叔	

頁碼	檔案原文摘錄	原書字形	備註
p. 964	●禍消	奠	
p. 981	●時第六十九	言 厶 間	
p. 981	●時篇	言 厶 間	
p. 988	葬避九空、地●	冎	
p. 1007	之而龜●	兆 厶 昔	
p. 1007	龜●	兆 厶 昔	
p. 1045	夫舩車●	形 食 人 方 口	
p. 1066	毋使其首●焉	卩 厶 冎	改為陷
p. 1205	言溶(潏)●而泉出	汜 厶 窟	