

# 《陳書》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  
2008/11/05 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 874 的「𠄎」字，Unicode 編碼是 4350，由於 XP 並不支援，仍需使用構字式「𠄎𠄎干」。

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：

《陳書》使用舊版造字 190 個，字頻 1624 次，這 190 個造字中，156 個可轉成 Windows XP 能顯示的字，字頻 1489 次；另外 34 個字必須轉成構字式，字頻 135 次。

轉碼完成製作轉碼分析表，請參考附件一《陳書》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 4930「岑」字，而原書使用之字為「岑」，查詢教育部異體字字典確認「岑」並非「岑」之異體字，因此歸為錯字，以程式取代為正確字形「岑」。
  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 3772 的「𠄎<sup>夕</sup>且<sup>夕</sup>」。

附件一、《陳書》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
799	𪗇	1	8E4D	9F5A	𪗇		
800	𪗈	1	8E4E	905D	𪗈		
821	𪗉	2	8E63	7718	𪗉		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
834	堯	2	8E70	712D	堯		
874	𠂔	1	8EBA	4350		𠂔 𠂔千	
896	櫛	1	8ED0	6AE9	櫛		
899	駘	3	8ED3	99DE	駘		
932	兌	1	8EF4	7683	兌		
938	燮	1	8EFA	7215	燮		
981	豺	2	8F66	72B2	豺		
982	胷	1	8F67	80F7	胷		胸，異體字問題。
1008	玦	1	8FA3	739E	玦		
1024	虵	3	8FB3	8675	虵		
1053	傲	3	8FD0	6160	傲		
1056	阨	10	8FD3	9641	阨		
1067	碁	4	8FDE	68CA	碁		
1091	漶	3	8FF6	58CD	漶		
1133	瓌	2	9061	74CC	瓌		
1237	稍	2	90EB	77DF	稍		
1513	糧	6	92C5	7CAE	糧		糧，異體字問題。
1650	賭	1	93B1	4417		月 厶者	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1893	袞	15	9548	886E	袞		
1897	祕	1	954C	8890	祕		祕，異體字問題。
1944	簪	3	957B	8A67	簪		
2077	鞮	1	9663	8F4A	鞮		
2199	鉦	1	9740	925F	鉦		
2503		1	98F5			焱△風	飈，異體字問題。
2505	飜	1	98F7	98DC	飜		
3772		1	8143			形 <sup>夕</sup> 且	待造字。
3773	鳴	1	8144	4CDF		明△鳥	
3774	觶	1	8145	89FD	觶		
3787	冢	12	8152	51A2	冢		
3792	涼	2	8157	51C9	涼		涼，異體字問題。
3804	効	10	8163	52B9	効		
3806	勳	3	8165	52E0	勳		
3814	却	13	816D	5374	却		
3815	𡗗	1	816E	537A	𡗗		
3820	厦	1	8173	53A6	厦		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3824	叙	2	8177	53D9	叙		
3826	叶	6	8179	53F6	叶		
3830	咤	3	817D	54A4	咤		
3835	噍	3	81A4	564D	噍		
3840	坂	2	81A9	5742	坂		
3846	壻	2	81AF	58FB	壻		
3849	獎	7	81B2	596C	獎		獎，異體字問題。
3858	冤	6	81BB	5BC3	冤		冤，異體字問題。
3860	尠	1	81BD	5C20	尠		
3864	峯	3	81C1	5CEF	峯		
3865	瀟	1	81C2	24162		灬 ㄩ 瀟	
3868	卮	1	81C5	5DF5	卮		
3877	迺	2	81CE	5EFC	迺		
3885		1	81D6			車 ㄩ 蘭	待造字。
3891	慙	21	81DC	6159	慙		
3892	喉	2	81DD	3B0B		日 ㄩ 侯	
3893		1	81DE			亻 ㄩ 妄	待造字。
3894	筋	1	81DF	4225		竹 ㄩ 勛	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3895	淑	1	81E0	6E86	淑		
3898	櫟	1	81E3	237C4		木 ㄩ 櫟	
3903	箐	1	81E8	25BE4		竹 ㄩ 韋	
3904	墟	1	81E9	588C	墟		
3905	叢	1	81EA	3AAB		睿 ㄩ 支	
3913	憇	4	81F2	20324		形 亻 〇 千 心 口	
3915	斲	2	8459	6585	斲		
3918	檀	1	81F7	65DC	檀		
3919	氈	54	81F8	23C0B		爽 ㄩ 毛	
3924	菴	5	81FD	671E	菴		
3925	𠂔	4	81FE	3775		𠂔	
3926		1	8240			承 ㄩ 灬	待造字。
3931	椀	1	8245	6900	椀		
3935	槩	4	8249	3BA3	槩		Unicode 字型 呈現差異。
3937	檝	4	824B	6A9D	檝		
3943		1	8251			形 亻 〇 无 心 口	形 亻 〇 无 心 口，錯字，以 程式全部取 代。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3947	舡	1	8255	4461		舟𠂔又	
3952	舢	1	825A	4470		舟𠂔步	
3956	渚	2	825E	6F74	渚		
3966	烟	1	8268	70DF	烟		
3974	牀	3	8270	7240	牀		
3978	猪	1	8274	732A	猪		
3982	珺	1	8278	73F7	珺		
3987	穀	1	827D	7474	穀		
3996	疎	15	82A8	758E	疎		
4010	阜	8	82B6	7690	阜		
4013	盃	1	82B9	76CC	盃		
4015	眦	1	82BB	7726	眦		
4029	褱	1	82C9	799D	褱		
4035	穿	1	82CF	7A7D	穿		
4038	竈	1	82D2	7AC8	竈		
4040	竝	324	82D4	7ADD	竝		
4041	豎	5	82D5	7AEA	豎		
4052	簾	1	82E0	7C34	簾		Unicode 字型 呈現差異。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4055	粃	1	82E3	7C83	粃		
4059	糇	1	82E7	7CC7	糇		
4068	綫	1	82F0	7DAB	綫		
4077	罇	2	82F9	7F47	罇		
4078		8	82FA			已合併	异，異體字問題。
4080	羣	79	82FC	7FA3	羣		
4090	脚	1	8347	811A	脚		
4091	臯	5	8348	81EF	臯		
4092	致	1	8349	81F6	致		
4095	舩	1	834C	8229	舩		
4104	莅	6	8355	8385	莅		
4105	菓	3	8356	83D3	菓		
4109	葛	2	835A	84AD	葛		
4115	藁	2	8360	85C1	藁		
4131	劦	5	8370	8842	劦		
4132		328	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	90	8373	885E	衛		衛，異體字問題。



編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4137	袜	1	8376	889C	袜		
4139	袴	2	8378	88B4	袴		
4142	裛	10	837B	8943	裛		
4146	覩	10	83A1	89A9	覩		睹，異體字問題。
4154	譽	1	83A9	8AD0	譽		
4155	訶	4	83AA	8B0C	訶		
4156	諂	2	83AB	8B1F	諂		
4157	譎	1	83AC	8B3F	譎		
4169	趁	1	83B8	8DA6	趁		
4172	屨	1	83BB	8E77	屨		
4173	躄	1	83BC	8EAD	躄		
4184	迹	19	83C7	8FF9	迹		
4195	鉢	1	83D2	9262	鉢		
4203	隣	14	83DA	96A3	隣		鄰，異體字問題。
4220	穎	2	83EB	9834	穎		
4222	頰	1	83ED	9847	頰		
4225	顛	1	83F0	9866	顛		
4240	髯	3	8440	9AF4	髯		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4242	鬪	8	8442	9B2D	鬪		
4252	鰓	6	844C			魚△眾	鰓，異體字問題。
4253	鳧	1	844D	9CF8	鳧		
4256	鷄	2	8450	9DC4	鷄		雞，異體字問題。
4260	麋	1	8454	9E9E	麋		
4267	鼈	1	845B	9F08	鼈		
4278	亘	3	8466	4E98	亘		
4287	織	1	846F	7E8E	織		
4303	謗	1	84A1	470E		言△勞	
4436	劔	2	8567	40AE		劔△石	
4437	廩	1	8568	5EEA	廩		
4442	潛	1	856D	6F98	潛		
4448		1	8573			形△厂 墨△	待造字。
4455	劔	24	857A			僉△刀	
4460	勅	73	85A1	52D1	勅		
4502	俎	2	85CB	723C	俎		
4515	呪	2	85D8	546A	呪		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4522	瞞	5	85DF	3B13		日 厶 票	
4699	暄	1	86F3	6685	暄		
4723	壇	1	874C	58C3	壇		
4735	壠	2	8758	58E0	壠		
4738	睺	1	875B	777A	睺		
4752	厯	1	8769	61D5	厯		
4794	蓁	1	87B5	8448	蓁		
4801	斌	1	87BC	5A2C	斌		
4902	兗	38	8862	5156	兗		
4930	岑	1	887E	5CBA	岑		岑，錯字，以程式全部取代。
5019	畧	1	88F9			畧	
5034	弃	16	8949	5F03	弃		棄，異體字問題。
5074	尫	1	8971	5C2B	尫		
5108	惛	1	89B5	60FD	惛		
5111	旃	1	89B8	65BE	旃		
5115	颯	1	89BC	4B03		風 厶 立	
5116	侑	5	89BD	4F37	侑		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5121		1	89C2			𠄎 主 方 攵 ㇀	傲，錯字，以程式全部取代。
5126	暮	2	89C7	8B29	暮		
5194		1	8A4C			𠄎 扌 尙 且 ㇀	擅，異體字問題。
5225	邨	2	8A6B	460F		血 厶 卩	
5232	擊	1	8A72	6480	擊		擊，錯字，以程式全部取代。
5287	壙	1	8ACB	367B		土 厶 廛	
5289	慤	1	8ACD	6164	慤		
5298	昨	1	8AD6	65FF	昨		
5302	昏	3	8ADA	662C	昏		
5338	暎	11	8AFE	668E	暎		
5355	𡗗	1	8B50	20AE4		𡗗	
5360	廩	1	8B55	913D	廩		
5437	并	88	8BC4	5E77	并		
5439	栢	1	8BC6	6822	栢		
5459	缺	1	8BDA	21647		垂 厶 夬	
5466	倏	1	8BE1	5010	倏		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5542	鑠	3	8C6E	93C1	鑠		
5549	藁	1	8C75	85F3	藁		
5551	賈	5	8C77	8CF7	賈		
5559	欸	2	8CA1	6B35	欸		
5578	弛	3	8CB4	38AE		弓厶屯	
5585	寓	8	8CBB	3762		宀会禹	
5591	杞	3	8CC1	233CC		木厶巳	杞，異體字問題。
5655	嚙	1	8D42	56BB	嚙		
5670	暈	1	8D51	66C5	暈		
5697	煑	1	8D6C	7151	煑		
5700	翮	4	8D6F	447D		舟厶翮	
5704	氈	1	8D73	6C0A	氈		氈，異體字問題。
5706	敘	14	8D75	654D	敘		敘，異體字問題。
5808	賀	1	8DFD	27D4D		夂会貝	
5860	无	1	C6D3	65E0	无		