

《抱朴子外篇》舊版造字轉碼說明

中研院資訊所文獻處理實驗室

中研院史語所漢籍電子文獻工作小組

2008/1/28 陳建安 製作

一、《抱朴子外篇》一書(抱朴子外篇.xml)使用舊版造字 182 個，字頻 852 次，詳如附件一。這 182 個造字中，138 個可轉成 Windows XP 能顯示的字，字頻 665 次；另外 44 個字必須轉成構字式，字頻 187 次。

二、附件一的造字分析表說明如下：

甲、編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、造字：舊版造字

丙、字頻(txt)：造字在「.txt」文件的出現次數

丁、字頻(xml)：造字在「.xml」文件的出現次數

戊、Big5：造字的 Big5 碼

己、Unicode：造字所對應的 Unicode 碼

庚、WinXP：造字在 Windows XP 的對應字形

辛、備註凡例：

1、校對問題，舊版漢籍錯字，可用程式全部取代：在舊版漢籍電子文獻中即存在的錯字，因校對時的疏漏而未更正，持續留存在新版漢籍電子文獻中；若該造字的所有頻次，皆屬於錯誤使用的錯字情形，可以用程式全部取代為正確字形。如編號 819 的「𠂔 𠂔侯」字，原字為「𠂔 𠂔侯」。

2、異體字問題：新版漢籍考量到使用者檢索及使用時的便利性，將用字原則改為除專詞等特殊情形之外，一律改用標準字呈現。如編號 4132 的「血 𠂔 𠂔」係「眾」字之異體，故以「眾」字取代。

3、Unicode 字型呈現差異：Unicode 字型與舊漢籍造字有些微差異，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「𠂔」字，Unicode 字型呈現為「𠂔」，實際上仍為同一字。

三、Unicode 目前收錄的漢字總數為 70194，分屬於三個不同區段，詳如表一。Windows XP 只支援 CJK 認同表意文字區的 20902 個字，內碼為 4E00-9FFF。所以如造字編號 874 的「𠂔」字，Unicode 碼為 4350，Windows XP 並不支援，故仍須使用構字式「𠂔 𠂔干」。

表一、Unicode 的字數及編碼區段

Unicode	字集子集合	新增字數	新增編碼區段	總字數	WinXP
1.1 版	CJK 認同表意文字區	20902	4E00-9FFF	20902	支援
3.0 版	CJK 認同表意文字擴充 A 區	6582	3400-4DFF	27484	不支援
3.1 版	CJK 認同表意文字擴充 B 區	42710	20000-2A6D6	70194	不支援

附件一、《抱朴子外篇》造字分析表

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
296	𠂔	7	7	FBEC			𠂔	
297	𠂕	19	19	FBED			𠂕	
298	𠂖	1	1	FBEE			𠂖	
299	𠂗	2	2	FBEF			𠂗	
300	𠂘	2	2	FBF0			𠂘	
301	𠂙	2	2	FBF1			𠂙	
549	𠂚	1	1	FDAF			𠂚	
568	𠂛	1	1	FDC2			𠂛	
658	𠂜	1	1	FE5D			𠂜	
707	𠂝	1	1	FEB0			𠂝	
717	𠂞	4	4	FEBA			𠂞	
730	𠂟	1	1	FEC7			𠂟	
750	𠂠	1	1	FEDB			𠂠	
758	𠂡	1	1	FEE3			𠂡	

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
766	𢇛	2	2	FEED			𢇛	
819	候	2	2	8E61	202EB		亻𡗗候	亻𡗗侯，校對問題，舊版漢籍錯字，可用程式全部取代。
843	鷺	1	1	8E79	9A61	鷺		
852	猓	1	1	8EA4	72E2	猓		
874	𡗗	1	1	8EBA	4350		𠂇𡗗千	
881	颺	15	15	8EC1	98C7	颺		
893	𡗗	5	5	8ECD	5CE9	𡗗		
900	𡗗	1	1	8ED4	5D73	𡗗		
981	豺	1	1	8F66	72B2	豺		
982	𡗗	3	3	8F67	80F7	𡗗		
1024	𡗗	3	3	8FB3	8675	𡗗		
1044	𡗗	3	3	8FC7	511B	𡗗		
1067	𡗗	8	8	8FDE	68CA	𡗗		
1097	𡗗	1	1	8FFC	37ED		山𡗗昏	
1205	𡗗	1	1	90CB	770E	𡗗		
1237	稍	1	1	90EB	77DF	稍		
1238	𡗗	1	1	90EC	77E6	𡗗		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
1418	𠂔	1	1	9244	4FF2	𠂔		
1478	𦰩	1	1	92A2			𦰩 𠂔 𠂔 𠂔	
1555	緡	1	1	92EF	7DCD	緡		
1565	緜	1	1	92F9	7DDC	緜		
1893	袞	7	7	9548	886E	袞		
1981	𦰪	4	4	95C2	8B90	𦰪		
2042	𦰫	2	2	9640	47FD		𦰫 𠂔 𠂔 𠂔	
2044	𦰬	6	6	9642	8E08	𦰬		
2152	𦰭	1	1	96D0	979F	𦰭		
2162	𦰮	2	2	96DA	918E	𦰮		
2167	𦰯	2	2	96DF	56D0	𦰯		
2502	𦰰	2	2	98F4	98C8	𦰰		
2527	𦰱	1	1	994E	22730		𦰱 𠂔 𠂔 𠂔	
2816	𦰲	1	1	9AF4	50C3	𦰲		
2947	𦰳	1	1	9BDA	8066	𦰳		
3021	𦰴	1	1	9C65			𦰴	
3063	𦰵	2	2	9CB1	2A51B		𦰵 𠂔 𠂔 𠂔	
3804	効	10	10	8163	52B9	効		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
3814	却	15	15	816D	5374	却		
3820	厦	5	5	8173	53A6	厦		
3830	咤	1	1	817D	54A4	咤		
3835	噍	2	2	81A4	564D	噍		
3836	喻	1	1	81A5	564F	喻		
3840	坂	1	1	81A9	5742	坂		
3845	躑	4	4	81AE	8E6E	躑		
3846	壻	8	8	81AF	58FB	壻		Unicode 呈現差異。
3849	獎	1	1	81B2	596C	獎		
3859	尔	1	1	81BC	5C14	尔		
3860	尠	14	14	81BD	5C20	尠		
3864	峯	3	3	81C1	5CEF	峯		
3873		1	1	81CA			广△段	廢，異體字問題。
3883	徧	1	1	81D4	5FA7	徧		
3891	慙	1	1	81DC	6159	慙		
3900	拞	1	1	81E5	62D5	拞		
3910	擘	2	2	81EF	64E5	擘		
3921	皙	1	1	81FA	6673	皙		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
3930	惇	2	2	8244	6901	惇		
3935	槩	5	5	8249	69E9	槩		Unicode 呈現差異。
3937	檝	2	2	824B	6A9D	檝		
3939	槩	1	1	824D	6AFD	槩		
3949	泣	2	2	8257	6D96	泣		
3964	威	1	1	8266	70D5	威		
3966	烟	5	5	8268	70DF	烟		
3974	牀	2	2	8270	7240	牀		
3976	犁	2	2	8272	7282	犁		
3980	獼	1	1	8276	7328	獼		
3981	珉	2	2	8277	73C9	珉		
3985	璫	1	1	827B	7447	璫		Unicode 呈現差異。
3996	疎	1	1	82A8	758E	疎		
4013	盃	1	1	82B9	76CC	盃		
4014	蓋	51	51	82BA	8462	蓋		
4015	眦	1	1	82BB	7726	眦		
4016	着	8	8	82BC	7740	着		Unicode 呈現差異。
4028	稟	6	6	82C8	7980	稟		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4035	穽	7	7	82CF	7A7D	穽		
4040	竝	45	45	82D4	7ADD	竝		
4041	豎	1	1	82D5	7AEA	豎		
4045	箒	1	1	82D9	7B92	箒		
4055	粦	1	1	82E3	7C83	粦		
4080	羣	81	81	82FC	7FA3	羣		Unicode 呈現差異。
4087	鰕	1	1	8344	4C47		魚𩺰旦	
4090	脚	11	11	8347	811A	脚		
4091	皐	4	4	8348	81EF	皐		
4103	茝	6	6	8354	831D	茝		
4104	莅	5	5	8355	8385	莅		
4107	葱	1	1	8358	8471	葱		
4109	葛	3	3	835A	84AD	葛		
4132		74	74	8371			血𩺰𩺰	眾，異體字問題。
4134	衛	10	10	8373	885E	衛		
4138	袞	1	1	8377	88A0	袞		
4139	袴	1	1	8378	88B4	袴		
4140	袵	3	3	8379	88B5	袵		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4142	裛	15	15	837B	8943	裛		
4146	覩	24	24	83A1	89A9	覩		
4158	譌	1	1	83AD	8B4C	譌		Unicode 呈現差異。
4165	獮	2	2	83B4	8C9B	獮		
4169	趁	1	1	83B8	8DA6	趁		
4172	屨	4	4	83BB	8E77	屨		Unicode 呈現差異。
4173	𪗇	1	1	83BC	8EAD	𪗇		
4184	迹	36	36	83C7	8FF9	迹		
4193	鈎	1	1	83D0	920E	鈎		
4198	鑛	4	4	83D5	945B	鑛		
4209	鞞	1	1	83E0	9789	鞞		
4226	飡	1	1	83F1	98E1	飡		
4233	𩇛	2	2	83F8	9A3B	𩇛		
4237	髡	2	2	83FC	9AE0	髡		
4242	鬪	7	7	8442	9B2D	鬪		Unicode 呈現差異。
4251	鰕	1	1	844B	9C15	鰕		
4253	鳬	1	1	844D	9CF8	鳬		
4254	鴟	1	1	844E	9D7E	鴟		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4256	鷄	5	5	8450	9DC4	鷄		
4265	龜	6	6	8459	9F03	龜		
4267	鼈	4	4	845B	9F08	鼈		
4437	廩	1	1	8568	5EEA	廩		
4455	劒	16	16	857A			僉 𠄎 刀	
4460	勅	1	1	85A1	52D1	勅		
4507	吳	11	11	85D0	5434	吳		
4530	咬	1	1	85E7	6648	咬		
4575	縊	3	3	8655	7E7F	縊		
4607	狔	2	2	8675	3E60		犴 𠄎 屯	
4723	塶	5	5	874C	58C3	塶		
4742		3	3	875F			衣 𠄎 執	褻，異體字問題。
4762	旃	1	1	8773	3ACB		𠄎 𠄎 冉	
4768	妬	4	4	8779	59AC	妬		
4816	𠄎	1	1	87CB			𠄎	
4817	寶	6	6	87CC	5BF3	寶		
4878	宀	2	2	884A	5B82	宀		Unicode 呈現差異。
4904	尅	3	3	8864	5C05	尅		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4985	瞢	1	1	88D7	669C	瞢		
4996	斲	1	1	88E2	65B5	斲		
5019	畧	5	5	88F9			畧	
5022	廌	1	1	88FC	5EBD	廌		
5034	弃	10	10	8949	5F03	弃		Unicode 呈現差異。
5043	彊	3	3	8952	5F4D	彊		
5046	彝	2	2	8955	5F5C	彝		
5065	忼	1	1	8968	5FFC	忼		
5075	𠂇	1	1	8972			𠂇	
5109	櫛	1	1	89B6	6A10	櫛		
5121		1	1	89C2			𠂇 丰 方 攴 𠂇	傲，校對問題，舊版漢籍錯字，可用程式全部取代。
5131	窻	1	1	89CC	7ABB	窻		Unicode 呈現差異。
5141	昊	2	2	89D6	23154		昊	
5169	𠂇	1	1	89F2	6239	𠂇		
5176		1	1	89F9			𠂇 𠂇 𠂇	沉，異體字問題。
5192	𠂇	4	4	8A4A	969F	𠂇		
5218		7	7	8A64			𠂇 𠂇 𠂇、𠂇	寬，異體字問題。

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
5225	邨	1	1	8A6B	460F		血𠂇卩	
5302	昏	1	1	8ADA	662C	昏		
5306	簡	7	7	8ADE			𦰩𠂇閒	
5310	烏	2	2	8AE2	8204	烏		
5328		1	1	8AF4			目𠂇𠂇	眇，校對問題，舊版漢籍錯字，可用程式全部取代。
5335	𧈧	2	2	8AFB	459F		亡𠂇虫	
5355	𧈧	1	1	8B50	20AE4		𧈧	
5371	𧈧	13	13	8B60	808E	𧈧		
5381	𧈧	2	2	8B6A	8103	𧈧		Unicode 呈現差異。
5396	鎔	1	1	8B79	938B	鎔		
5450	𧈧	4	4	8BD1	6E4C	𧈧		
5466	𧈧	3	3	8BE1	5010	𧈧		Unicode 呈現差異。
5537	𧈧	2	2	8C69	79BC	𧈧		
5556	隸	1	1	8C7C	96B7	隸		
5562	𧈧	1	1	8CA4	7A09	𧈧		
5578	𧈧	1	1	8CB4	38AE		弓𠂇𧈧	
5580	𧈧	1	1	8CB6	79BC	𧈧		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
5697	𪗇	1	1	8D6C	7151	𪗇		Unicode 呈現差異。
5702	𪗈	1	1	8D71	53A8	𪗈		
5704	𪗊	1	1	8D73	6C0A	𪗊		
5706	𪗌	16	16	8D75	654D	𪗌		
5709	𪗏	3	3	8D78	872F	𪗏		
5724	𪗔	2	2	8DA9	7A38	𪗔		
5753	𪗙	2	2	8DC6	6F5C	𪗙		
5768	𪗜	5	5	8DD5	6FF6	𪗜		
5808	𪗠	1	1	8DFD	27D4D		𪗠𪗡	