

《二刻拍案驚奇》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
2008/8/13 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 74 的「髻」字，Unicode 編碼是 29B06，由於 XP 並不支援，仍需使用構字式「髻𠄎𠄎」。

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：

《二刻拍案驚奇》使用舊版造字 158 個，字頻 2980 次，這 158 個造字中，148 個可轉成 Windows XP 能顯示的字，字頻 2939 次；另外 9 個字必須轉成構字式，字頻 17 次，1 字轉成標記，字頻 24 次。

轉碼完成製作轉碼分析表，請參考附件一《二刻拍案驚奇》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. 無此字形，以標記符號手動修改：原書無此字形，但該處在書中字形大小不同，因此需把舊版造字手動修改為語言所慣用標記符號。例：檔案中使用造字編號 4171 的「足厂封 - 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 4263 的「麩」字，Unicode 字型呈現為「麩」，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 3097 的「彳幸」。

附件一、《二刻拍案驚奇》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
20	藪	4	FA53			𣎵𣎵救	
74	髻	1	FAAB	29B06		髻𣎵𣎵	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
893	峩	1	8ECD	5CE9	峩		
912	鷓	1	8EE0	9DC0	鷓		
1490	籛	1	92AE	7C5D	籛		
1491	擷	4	92AF	6527	擷		
1513	糧	3	92C5	7CAE	糧		糧，異體字問題。
1547	𠂇	1	92E7	4E31	𠂇		
1616	逕	4	936D	284E6		迨△里	
1691	髡	1	93DA	4BFC		髡△狄	
1875	割	1	94F5	5268	割		
1977	奮	1	95BE	5964	奮		
1981	讐	3	95C2	8B90	讐		
2046	踪	44	9644	8E2A	踪		
2053	躡	1	964B	8EA7	躡		
2109	簞	1	96A5	4261		𦰃△簞	
2154	魀	4	96D2	9B40	魀		
2157	酌	7	96D5	9167	酌		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2409	闕	1	9875	95A7	闕		
2505	翻	3	98F7	98DC	翻		
2592	臟	5	99B1	8D1C	臟		
2999	嗷	5	9C4F	5523	嗷		
3016	闕	1	9C60	28D91		門△臭	
3097		3	9CD3			彳△幸	待造字。
3098	揼	1	9CD4	63F8	揼		Unicode 字型 呈現差異。
3789	冲	8	8154	51B2	冲		冲，異體字問 題。
3791	况	23	8156	51B5	况		况，異體字問 題。
3792	凉	9	8157	51C9	凉		凉，異體字問 題。
3794	湊	36	8159	51D1	湊		湊，異體字問 題。
3796	决	3	815B	51B3	决		决，異體字問 題。
3801	剗	4	8160	5257	剗		
3804	効	18	8163	52B9	効		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3805	勅	1	8164	52C5	勅		
3818	廩	1	8171	53AB	廩		
3819	廩	9	8172	53AE	廩		廩，異體字問題。
3820	厦	1	8173	53A6	厦		
3824	叙	4	8177	53D9	叙		
3825	疊	27	8178	53E0	疊		
3829	咏	1	817C	548F	咏		
3830	咤	9	817D	54A4	咤		
3834	嗶	1	81A3	5637	嗶		
3843	堦	5	81AC	5826	堦		
3846	堦	56	81AF	58FB	堦		
3850	姪	8	81B3	59D9	姪		
3851	姪	1	81B4	59F9	姪		
3852	媼	1	81B5	5A63	媼		
3855	孽	1	81B8	5B7C	孽		孽，異體字問題。
3858	冤	102	81BB	5BC3	冤		冤，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3864	峯	1	81C1	5CEF	峯		
3868	卮	14	81C5	5DF5	卮		
3876	廻	11	81CD	5EFB	廻		
3886	悞	1	81D7	609E	悞		
3891	慙	11	81DC	6159	慙		
3906	摧	1	81EB	6409	摧		
3909	擡	85	81EE	64E1	擡		
3921	皙	2	81FA	6673	皙		
3924	菴	1	81FD	671E	菴		期，異體字問題。
3931	椀	2	8245	6900	椀		
3934	榻	1	8248	69C5	榻		
3937	檝	1	824B	6A9D	檝		
3938	橈	2	824C	6AC8	橈		
3966	烟	13	8268	70DF	烟		
3974	牀	11	8270	7240	牀		
3976	犁	1	8272	7282	犁		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3978	猪	16	8274	732A	猪		豬，異體字問題。
3979	猫	2	8275	732B	猫		貓，異體字問題。
3994	畊	3	82A6	754A	畊		
3996	疎	14	82A8	758E	疎		
4000	疴	2	82AC	75B4	疴		
4010	臯	1	82B6	7690	臯		
4014	蓋	1	82BA	8462	蓋		
4016	着	1797	82BC	7740	着		
4028	稟	10	82C8	7980	稟		
4036	窰	2	82D0	7AB0	窰		
4038	竈	16	82D2	7AC8	竈		
4040	竝	2	82D4	7ADD	竝		
4042	笋	4	82D6	7B0B	笋		
4043	筴	2	82D7	7B53	筴		筴，異體字問題。
4044	筴	4	82D8	7B6F	筴		
4045	筴	3	82D9	7B92	筴		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4065	紮	12	82ED	7D25	紮		
4067	綉	7	82EF	7D89	綉		
4069	綑	2	82F1	7DB3	綑		
4077	罇	5	82F9	7F47	罇		
4080	羣	34	82FC	7FA3	羣		
4089	脇	1	8346	8107	脇		
4091	臯	46	8348	81EF	臯		
4105	菓	21	8356	83D3	菓		
4107	葱	2	8358	8471	葱		
4112	蔴	12	835D	8534	蔴		
4114	藥	1	835F	8616	藥		
4115	藁	2	8360	85C1	藁		
4134	衛	16	8373	885E	衛		衛，異體字問題。
4137	袜	1	8376	889C	袜		
4139	袴	3	8378	88B4	袴		
4144	霸	1	837D	8987	霸		Unicode 字型呈現差異。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4145	羈	4	837E	898A	羈		Unicode 字型呈現差異。
4146	覩	4	83A1	89A9	覩		睹，異體字問題。
4171		24	83BA			𠄎足厂 封	無此字形，以標記符號手動修改。
4172	屨	2	83BB	8E77	屨		
4173	𨇗	9	83BC	8EAD	𨇗		
4176	輒	1	83BF	8F19	輒		
4184	迹	4	83C7	8FF9	迹		
4193	鈎	8	83D0	920E	鈎		鈎，異體字問題。
4195	鉢	4	83D2	9262	鉢		
4226	飡	1	83F1	98E1	飡		
4232	隲	2	83F7	96B2	隲		
4242	鬪	3	8442	9B2D	鬪		
4256	鷄	21	8450	9DC4	鷄		雞，異體字問題。
4263	麩	1	8457	9EAF	麩		Unicode 字型呈現差異。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4267	鼈	1	845B	9F08	鼈		
4307	罵	38	84A5	99E1	罵		罵，異體字問題。
4308	劫	3	84A6	5227	劫		
4309	劫	4	84A7	523C	劫		
4432		3	8563			形糸厂 墨	纏，異體字問題。
4447	捏	16	8572	63D1	捏		
4460	勅	3	85A1	52D1	勅		
4512		2	85D5			口	咤，異體字問題。
4515	呪	1	85D8	546A	呪		
4518	响	1	85DB	54CD	响		
4742		3	875F			衣	褻，異體字問題。
4768	妬	10	8779	59AC	妬		妒，異體字問題。
4785	贗	2	87AC	8D0B	贗		
4805	瀆	6	87C0	51DF	瀆		瀆，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4903	𠄎	5	8863	5C00	𠄎		
4904	尅	2	8864	5C05	尅		
4907	橐	1	8867	69D6	橐		
4980	𤇗	1	88D2			𤇗 彳 火	
5034	弃	1	8949	5F03	弃		棄，異體字問題。
5174		1	89F7			𠄎 刀 兂 、 𠄎	兔，異體字問題。
5176		12	89F9			冫 𠄎 兂	沉，異體字問題。
5180	拘	1	89FD	6285	拘		
5229	撐	3	8A6F	6490	撐		
5240	擗	1	8A7A	6481	擗		
5290	携	11	8ACE	643A	携		
5299	戟	1	8AD7	39B8		卓 𠄎 戈	戟，異體字問題。
5300	鬪	9	8AD8	9B2A	鬪		
5347	憾	3	8B48	617D	憾		
5434	湟	1	8BC1	6E7C	湟		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5437	并	10	8BC4	5E77	并		
5439	栢	2	8BC6	6822	栢		
5464	煥	2	8BDF	7157	煥		
5478	飴	12	8BED	4D77	飴		
5551	賈	17	8C77	8CF7	賈		
5555	迥	1	8C7B	9008	迥		
5558	龕	3	8C7E	9E84	龕		
5582	鞣	1	8CB8	97B5	鞣		
5583	殮	2	8CB9	98F1	殮		
5588	恡	2	8CBE	6061	恡		
5590	朶	8	8CC0	6736	朶		
5591	杞	6	8CC1	233CC		木𠂇巳	杞，異體字問題。
5595	朶	5	8CC5	579C	朶		
5666	汚	4	8D4D	6C5A	汚		汚，異體字問題。
5702	厨	2	8D71	53A8	厨		厨，異體字問題。
5706	敘	29	8D75	654D	敘		敘，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
5742	澁	1	8DBB	6F81	澁		
5768	濶	4	8DD5	6FF6	濶		闊，異體字問題。
5795	媵	1	8DF0	3715		女媵	

二、手動取代表說明：

《二刻拍案驚奇》檔案中需手動修改的情況分為以下幾種：

(一) 《二刻拍案驚奇》檔案中之舊版造字，因字形需手動修改為語言所慣用標記，並製作手動取代表，請參考附件二。欄位說明如下：

1. 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
2. 造字：舊版造字字形。
3. 檔案原文摘錄：摘錄檔案《二刻拍案驚奇》中的文句段落，其中紅字底線的部分，為需要手動取代的原文。
4. 手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。
5. 備註：記錄修改相同辭句的次數或其他事項。

附件二、《二刻拍案驚奇》手動取代表

編號	造字	檔案原文摘錄	手動取代結果	備註
4171	𠃉𠃊𠃋𠃌	軒轅台 <u>森森森</u> 乃黃帝 鑄鼎於此。 <u>森森森</u>	軒轅台☆乃黃帝鑄鼎 於此。★	

編號	造字	檔案原文摘錄	手動取代結果	備註
		酒香亭 <u>森森森</u> 乃漢武帝得仙酒於此。 <u>森森森</u>	酒香亭 ☆ 乃漢武帝得仙酒於此。★	
		朗吟亭 <u>森森森</u> 乃呂仙遺蹟。 <u>森森森</u>	朗吟亭 ☆ 乃呂仙遺蹟。★	
		柳毅井 <u>森森森</u> 乃柳毅為洞庭龍女傳書處。 <u>森森森</u>	柳毅井 ☆ 乃柳毅為洞庭龍女傳書處。★	

(二)《二刻拍案驚奇》檔案中出現「●」字形，共計 117 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，請參考附件三欄位說明如下：

1. 頁碼：「●」字形所在位置之原書頁碼。
2. 檔案原文摘錄：摘錄《二刻拍案驚奇》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。

附件三、《二刻拍案驚奇》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 8	●●喝喝	口 厶 口 厶	有 4 處
p. 8	那裏●	扌 牢	
p. 24	手段高得●岬	山 宀 犬	改為山 突
p. 24	在旁邊●着觀看	躑	

頁碼	檔案原文摘錄	原書字形	備註
p. 27	點指●腳教人	扌△畫	
p. 27	只做不●	係	
p. 28	小道人●了出來	扌△雙	
p. 42	今日●得賠還他	扌△棄	
p. 58	●脫了身子	扌△劣	待造字
p. 68	萬一●驚起來	扌	
p. 68	●寶金壺一把	筴	
p. 71	見說家主要●	闕	
p. 71	料他是個要●	闕	
p. 71	只是眼●	闕	
p. 73	取得來送與●	行△亢	
p. 73	●人家	衍	
p. 74	到別處去●	扌△牢	
p. 77	只要●老的	闕	有 2 處
p. 78	還要●	闕	
p. 79	在湯家●的正是	闕	
p. 95	一齊●了家來	逵	
p. 114	有些●魀	鬼△監	

頁碼	檔案原文摘錄	原書字形	備註
P. 135	沙邊瀾●	涑△鳥	
P. 144	為●了廝鬧	闕	
P. 149	●紅坐六的	𠂔扌夫手𠂔	改為捧
p. 157	向桌上一●道	扌△雙	
p. 160	●梅香認合	彳△芻	
p. 165	我不●他	係	
p. 165	鳳生覷●	覷	
p. 171	不要●他	係	
p. 187	●尪起來	鬼△監	
p. 208	別處●酒	噏	
p. 209	有些不●尪	鬼△監	
p. 217	擁●他去睡了	扌△族	
p. 225	上等的●	行△亢	
p. 225	●，臺州數一數二	行△完	
p. 235	帶胎●鹿	麕	
p. 236	●中有	厨	改為廚
p. 266	只把●子拴了	孛△頁	
p. 267	剩得不少，●	行△亢	

頁碼	檔案原文摘錄	原書字形	備註
p. 267	●家本色	行△完	
p. 268	弄得不●●	𠵼 兀△介	改為𠵼𦉳
p. 271	婦人不好●得	扌	
p. 280	就便●不轉來	扌	
p. 282	略把●子一伸	孛△頁	
p. 321	何規●不打	言△摹	待造字
p. 333	●砒	步△匕	改為死
p. 333	●汞	步△匕	改為死
p. 337	不●●	𠵼 兀△介	改為𠵼𦉳
P. 338	●呵力重	口△么	
p. 338	洗●伐毛	髓	
p. 351	不敢違●	扌	
p. 354	綠●袞冕	韋△支	
p. 355	怒●●不息	口△牛口△牛	待造字
p. 356	老大一片跣●	足△荅	
p. 367	有些●𦉳	鬼△監	

頁碼	檔案原文摘錄	原書字形	備註
p. 377	光景有些●●	𧨇 兀△介	改為𧨇𧨇
p. 378	●宿情濃	闕	
p. 383	睡得惺●	惚	
p. 388	●喝一聲	口△么	
p. 402	北人就去戒●	厨	改為廚
p. 404	我們●着鎗棒	拏	
P. 429	●賭行中	闕	
p. 429	非賭即●	闕	
P. 434	●盤上螞蟻	金△教	
p. 464	心性古●	忖△敝	
P. 484	名曰黃圻●	嶠	
P. 488	●在李方哥	扌△塞	
p. 493	夜轉來。●了	扌△擻	
p. 501	着朋友●	行△亢	
p. 501	●人家走動兩番	行△完	
p. 504	哼哼●●	口△賴口△賴	
p. 508	花枝●孃	女△梟	

頁碼	檔案原文摘錄	原書字形	備註
p. 521	香爐裏亂●	扌厶塞	
p. 521	●罷	扌厶塞	
p. 527	料●他不倒	扌厶班	
p. 549	大●恩典	賚	
p. 594	只得●曲作直	扌	
p. 605	買祠部度●	牒	改為牒
p. 611	亂●亂用	扌厶塞	
p. 618	慌忙●了進來	遼	
p. 622	放下了跣●	足厶荅	
p. 624	途中遭雨濕●	形广秝土口	待造字
p. 624	●多發了斑點	黑厶彡	
p. 630	一個人來●他	係	
p. 635	靦●着面龐	覷	
p. 635	勾着●子	孛厶頁	
p. 638	●到楊二郎	遼	
P. 640	看來是個●	行厶亢	
p. 640	●人家了	行厶完	
p. 649	這一番●●	嗶噉	

頁碼	檔案原文摘錄	原書字形	備註
p. 649	形如蠶●	范△虫	
p. 652	盡●將來	逵	
p. 657	見他夫婦惺●	惚	
p. 662	曬●那帽子時	日△良	
p. 662	●着面且認	覘	
p. 664	髮髻●在箱內	扌△塞	
p. 676	●一味懷忠	衡	
p. 688	燕小乙哥走入●	行△亢	
p. 688	●人家吃酒去了	行△完	
p. 688	風銷焰●	爝	
p. 689	做箇不惺●	惚	
p. 689	何處兒郎真●懂	夢	