



# 《繡像金瓶梅詞話》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  
2010/01/18 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 14 的「胤」字，Unicode 編碼是 38E7，由於 XP 並不支援，仍需使用構字式「彳么月廾」。

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：

《繡像金瓶梅詞話》使用舊版造字 292 個，字頻 7430 次，這 292 個造字中，224 個可轉成 Windows XP 能顯示的字，字頻 6426 次；另外 68 個字必須轉成構字式，字頻 1004 次。

轉碼完成製作轉碼分析表，請參考附件一《繡像金瓶梅詞話》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. **部份使用不同字形，手動修改**：檔案中為同一舊版造字，校對時卻發現原書使用兩個以上不同的字形，可能部份舊版造字為錯字或原書使用兩種異體字形，導致同一舊版造字轉碼時對應到兩個以上的字形；將一部份手動修改，剩餘的頻次則用程式取代成對應的 Unicode 或構字式。例：造字編號 2206 的「祆」字頻次為 5，依書將 1 字手動修改為「祆」，剩餘 4 字轉為「祆」。手動修改部份請參考附件二《繡像金瓶梅詞話》手動取代表。
3. **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 771 的「攵」字，而原書使用之字為「攵」，查詢教育部異體字字典確認「攵」並非「攵」之異體字，因此歸為錯字，以程式取代為正確字形「攵」。
4. **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。
5. **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 3893 的「亻𠂔妄」。
6. **依原書修改未定義字形**：編號 309 舊版造字「㊦」代表未定義字

型，依原書修改為書中字形。例：「𠄎」在書中的字形為「𠄎」，依原書修改為「𠄎」。

附件一、《繡像金瓶梅詞話》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
14	胤	1	FA4D	38E7		𠄎 彳 幺 月 乚 口	
20	籟	4	FA53			𠄎 叕 救	
74	髻	5	FAAB	29B06		髻 𠄎 丐	
296	𠄎	125	FBEC			𠄎	
297	𠄎	394	FBED			𠄎	
298	𠄎	76	FBEE			𠄎	
299	𠄎	26	FBEF			𠄎	
300	𠄎	26	FBF0			𠄎	
308	∞	1	FBF8			∞	
309	𠄎	1	FBF9			𠄎	𠄎，依原書修改未定義字形。
426	𠄎	1	FCD1			𠄎	
434	𠄎	1	FCD9			𠄎	
521	𠄎	1	FD71			𠄎	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
639	彳	1	FE4A			彳	
658	疒	7	FE5D	7592		疒	
675	食	8	FE6E			食	
690	𠂇	1	FE7D			𠂇	
707	ㄍ	4	FEB0			ㄍ	
714	ト	1	FEB7			ト	
717	疋	12	FEBA			疋	
726	丿	1	FEC3			丿	向，錯字，以程式全部取代。
742	ㄥ	1	FED3			ㄥ	
750	丩	3	FEDB			丩	
752	𠂇	7	FEDD			𠂇	
761	久	1	FEE6			久	
770	扌	102	FEEF	624C		扌	
771	攵	1	FEF0			攵	久，錯字，以程式全部取代。
777	宀	1	FEF6			宀	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
781	丿	1	FEFA			丿	
814	墓	2	8E5C	87C7	墓		
841	礪	4	8E77	7932	礪		
852	貉	1	8EA4	72E2	貉		
893	峩	1	8ECD	5CE9	峩		
900	峩	1	8ED4	5D73	峩		
912	鷓	1	8EE0	9DC0	鷓		
938	燮	3	8EFA	7215	燮		
1014	玠	1	8FA9	73CE	玠		
1067	碁	2	8FDE	68CA	碁		
1100	身	7	9040	8EAB	身		
1151	劓	13	9073	5273	劓		
1170	瞳	1	90A8	7583	瞳		
1187	盖	11	90B9	76D6	盖		
1201	昵	3	90C7	7724	昵		
1213	睽	15	90D3	7743	睽		
1268		1	914B			灬 𠃉 鑿	待造字。
1278	碍	16	9155	788D	碍		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1310	蠟	3	9175	45EB		虫 𧈧 麻	
1321	確	2	91A2	7936	確		
1400	窓	63	91F1	7A93	窓		
1467	纂	5	9275	25C67		𦰃 𠂇 秦	
1491	擷	3	92AF	6527	擷		
1513	糧	16	92C5	7CAE	糧		糧，異體字問題。
1598	翰	2	935B	4A7A		革 𧈧 翁	
1616	運	4	936D	284E6		辶 𠂇 里	
1641	胆	20	93A8	80C6	胆		膽，異體字問題。
1642	脉	57	93A9	8109	脉		
1646	脗	2	93AD	8117	脗		
1668	臍	4	93C3	81C8	臍		
1691	髻	58	93DA	4BFC		髻 𠂇 狄	
1703	璇	8	93E6	2773D		衤 𧈧 旋	
1781	葢	13	9475	854B	葢		
1893	袞	6	9548	886E	袞		
1905	裊	6	9554	88E9	裊		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1912	襍	9	955B	894D	襍		
1926	覓	3	9569	8994	覓		
1981	讐	5	95C2	8B90	讐		
1994	達	1	95CF	6FBE	達		
2021	趁	1	95EA	8D82	趁		
2042	䟽	1	9640	47FD		足厶充	
2044	踈	6	9642	8E08	踈		
2046	踪	15	9644	8E2A	踪		
2053	躡	27	964B	8EA7	躡		
2072		1	965E			𠄎厶縻	𠄎厶縻，錯字，以程式全部取代。
2151	𠄎	2	96CF	961D		𠄎	
2157	酌	8	96D5	9167	酌		
2162	醜	7	96DA	918E	醜		
2189	瞿	1	96F5	23C3D		毛厶瞿	
2190	稔	4	96F6	6BFA	稔		
2206	袂	5	9747	8884	袂		部份使用不同字形「袂」，手

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							動修改。
2409	閔	2	9875	95A7	閔		
2432	鞞	2	98AE	29350		革厶登	
2449	雒	1	98BF			名厶隹	
2452	蠳	1	98C2	45F6		虫厶葛	
2459	騷	1	98C9	4BAB		馬厶婁	
2502	颺	1	98F4	98C8	颺		
2505	翻	1	98F7	98DC	翻		
2758	猷	5	9ABA	732E	猷		猷，異體字問題。
2764	窃	1	9AC0	7A83	窃		
2770	絳	1	9AC6	7D98	絳		
2885	叫	93	9B7A	544C	叫		
2893	属	1	9BA4	5C5E	属		
2896	庙	7	9BA7	5E99	庙		廟，異體字問題。
2905	挀	1	9BB0	6455	挀		
2937	豎	1	9BD0			𠄎厶立	
2975	邈	1	9BF6	8FEF	邈		



編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2997	盐	6	9C4D	25081		形土人 口皿	
2999	咆	7	9C4F	5523	咆		
3004	籊	1	9C54	7BD0	籊		
3020	忝	3	9C64			忝	
3021	忝	10	9C65	5FC4		忝	
3029	鋤	5	9C6D	941D	鋤		
3096	歲	1	9CD2	5D57	歲		
3098	揼	9	9CD4	63F8	揼		Unicode 字型 呈現差異。
3113	彡	1	9CE3	6245	彡		
3160	坟	33	9D53	575F	坟		
3161	壘	22	9D54	58DC	壘		
3164		1	9D57			土册	土册，異體 字問題。
3177	蟻	3	9D64	7F4E	蟻		
3230	芦	3	9DBB	82A6	芦		
3250	挑	1	9DCF	6304	挑		
3251	灯	10	9DD0	706F	灯		燈，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
3770	个	34	8141	4E2A	个		
3789	冲	5	8154	51B2	冲		沖，異體字問題。
3791	况	4	8156	51B5	况		況，異體字問題。
3794	凑	6	8159	51D1	凑		湊，異體字問題。
3797	凭	5	815C	51ED	凭		
3804	効	9	8163	52B9	効		
3805	勅	16	8164	52C5	勅		
3807	驗	1	8166	9A10	驗		
3814	却	260	816D	5374	却		
3815	𡗗	2	816E	537A	𡗗		
3819	廝	34	8172	53AE	廝		廝，異體字問題。
3824	叙	9	8177	53D9	叙		
3825	疊	16	8178	53E0	疊		
3829	咏	3	817C	548F	咏		





編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3834	嗶	2	81A3	5637	嗶		
3836	喻	1	81A5	564F	喻		
3843	堦	44	81AC	5826	堦		
3850	姪	2	81B3	59D9	姪		
3852	媼	1	81B5	5A63	媼		
3853	褻	12	81B6	2776C		敝衤衣	
3858	寃	13	81BB	5BC3	寃		冤，異體字問題。
3864	峯	10	81C1	5CEF	峯		
3868	卮	2	81C5	5DF5	卮		
3870	帑	87	81C7	5E0B	帑		
3876	廻	7	81CD	5EFB	廻		
3877	迺	1	81CE	5EFC	迺		
3886	悞	36	81D7	609E	悞		
3891	慙	1	81DC	6159	慙		
3893		4	81DE			亻 㒼 妄	待造字。
3899	担	10	81E4	62C5	担		
3909	擡	126	81EE	64E1	擡		
3912	拈	1	81F1	62C3	拈		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3924	朥	1	81FD	671E	朥		
3927	柏	1	8241	6855	柏		
3934	榻	36	8248	69C5	榻		
3935	槩	6	8249	3BA3	槩		Unicode 字型 呈現差異。
3938	橈	41	824C	6AC8	橈		
3940	欝	6	824E	6B1D	欝		
3942	毡	74	8250	6BE1	毡		
3944	冰	14	8252	6C37	冰		
3951	湏	3	8259	6E4F	湏		
3961	灾	8	8263	707E	灾		
3963	炁	12	8265	7081	炁		
3966	烟	92	8268	70DF	烟		
3968	煨	1	826A	7145	煨		
3974	牀	127	8270	7240	牀		
3975	牕	18	8271	7255	牕		Unicode 字型 呈現差異。
3976	犁	1	8272	7282	犁		
3977	狗	1	8273	72E5	狗		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3978	猪	38	8274	732A	猪		Unicode 字型 呈現差異。
3979	猫	2	8275	732B	猫		
3984	瑠	3	827A	7460	瑠		
3990	甄	1	82A2	750E	甄		
3996	疎	17	82A8	758E	疎		
4001	瘞	1	82AD	7602	瘞		
4016	着	215 8	82BC	7740	着		
4017	螻	5	82BD	880F	螻		
4025	礪	3	82C5	7934	礪		
4028	稟	1	82C8	7980	稟		
4034	穉	2	82CE	7A49	穉		
4036	窰	4	82D0	7AB0	窰		
4038	竈	29	82D2	7AC8	竈		
4041	豎	4	82D5	7AEA	豎		
4042	笋	24	82D6	7B0B	笋		
4044	筋	60	82D8	7B6F	筋		
4067	綉	189	82EF	7D89	綉		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4069	綑	4	82F1	7DB3	綑		
4074	縑	23	82F6	7E27	縑		
4080	羣	3	82FC	7FA3	羣		
4085	耻	23	8342	803B	耻		
4089	脇	1	8346	8107	脇		
4090	脚	195	8347	811A	脚		
4091	臯	3	8348	81EF	臯		
4105	菓	175	8356	83D3	菓		
4107	葱	17	8358	8471	葱		
4112	蔴	13	835D	8534	蔴		
4129	虬	1	836E	866C	虬		
4132		714	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	5	8373	885E	衛		衛，異體字問題。
4139	袴	7	8378	88B4	袴		
4140	衽	13	8379	88B5	衽		
4146	覩	6	83A1	89A9	覩		
4167	賈	4	83B6	8CEB	賈		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4171		1	83BA			𠂇足厂 封	躑，異體字問題。
4173	𠂇	26	83BC	8EAD	𠂇		
4176	輒	6	83BF	8F19	輒		
4184	迹	3	83C7	8FF9	迹		
4193	鈎	33	83D0	920E	鈎		
4203	隣	4	83DA	96A3	隣		鄰，異體字問題。
4218	菲	15	83E9	97EE	菲		
4226	飡	14	83F1	98E1	飡		
4232	隲	6	83F7	96B2	隲		
4237	髡	3	83FC	9AE0	髡		部份使用不同字形「髡𠂇八」，手動修改。
4242	鬪	5	8442	9B2D	鬪		
4247	鮎	1	8447	9B8E	鮎		
4252	鰓	4	844C			魚𠂇眾	鰓，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4256	鷄	25	8450	9DC4	鷄		
4262	麪	4	8456	9EAA	麪		Unicode 字型 呈現差異。
4263	麩	1	8457	9EAF	麩		Unicode 字型 呈現差異。
4278	亘	3	8466	4E98	亘		
4279	驄	2	8467	9A18	驄		
4329	倨	1	84BB	4FB0	倨		
4400	兇	4	8543	5160	兇		
4447	捏	61	8572	63D1	捏		
4455	劒	3	857A			僉  刃	
4460	勅	4	85A1	52D1	勅		
4461	勗	5	85A2	52D4	勗		
4463	摧	1	85A4	3A72		扌  崔	
4477		1	85B2			 艹 勺 一 	芍，異體字問 題。
4496	段	145	85C5	53DA	段		
4515	呪	11	85D8	546A	呪		
4518	响	44	85DB	54CD	响		



編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4520	邈	1	85DD	285FF		辶△豸	
4560	協	2	8646	604A	協		
4572	嘍	1	8652	20DFA		口△勹	
4616	壳	5	867E	58F3	壳		
4672		1	86D8			豸△畺	待造字。
4688		1	86E8			土△屏	待造字。
4715	嘍	1	8744	5688	嘍		
4718	屯	1	8747	340C		屯	
4750	卅	16	8767	5344	卅		卅，錯字，以程式全部取代。
4759	禰	2	8770	8947	禰		
4768	妬	7	8779	59AC	妬		
4779	鞞	2	87A6	97B2	鞞		
4784	鞞	1	87AB	9793	鞞		
4805	瀆	7	87C0	51DF	瀆		瀆，異體字問題。
4816	亻	8	87CB			亻	
4823	窑	1	87D2	7A91	窑		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4848	輩	2	87EB	8EF0	輩		
4891	類	3	8857	4AEB		婁厶頁	
4903	吋	8	8863	5C00	吋		
4904	尅	8	8864	5C05	尅		
4966	癩	1	88C4	764E	癩		
4981	鳧	2	88D3	9CEC	鳧		
5008	幪	1	88EE	5E5E	幪		
5013	汜	9	88F3			汜	
5032	焜	1	8947	3B0E		焜	
5075	犛	1	8972	72AD		犛	
5094	翯	2	89A7	26407		翯	
5101	馱	11	89AE	4B7E		馬厶犬	
5123	愧	1	89C4	6130	愧		
5147	恠	1	89DC	6060	恠		
5173	漱	2	89F6	6F44	漱		
5176		23	89F9			汜厶冗	沉，異體字問題。
5180	拘	20	89FD	6285	拘		
5185	柄	1	8A43	62A6	柄		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5188	搨	1	8A46	62EA	搨		
5229	撐	19	8A6F	6490	撐		
5230		2	8A70			形土畝 且	壇，異體字問題。
5240	摔	1	8A7A	6481	摔		
5241	鷺	3	8A7B	9D5E	鷺		
5245	擻	1	8AA1			掉	
5281	餓	10	8AC5	994A	餓		
5290	携	25	8ACE	643A	携		
5299	戟	4	8AD7	39B8		卓	戟，異體字問題。
5300	鬪	47	8AD8	9B2A	鬪		
5392	悤	1	8B75	60A4		悤	
5395	規	1	8B78	4093		矢	
5437	并	139	8BC4	5E77	并		
5439	栢	7	8BC6	6822	栢		
5440	昧	1	8BC7	771C	昧		
5447	躡	1	8BCE	8E83	躡		
5493	之	1	8BFC			之	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5543	蝥	1	8C6F	9B9D	蝥		
5551	費	1	8C77	8CF7	費		
5555	迥	1	8C7B	9008	迥		
5556	隸	5	8C7C	96B7	隸		
5558	龕	1	8C7E	9E84	龕		
5559	欸	10	8CA1	6B35	欸		
5583	殮	1	8CB9	98F1	殮		
5595	塚	6	8CC5	579C	塚		
5621	泪	14	8CDF	6CEA	泪		
5666	污	11	8D4D	6C5A	污		污，異體字問題。
5694	槩	1	8D69	3BA3	槩		
5702	厨	24	8D71	53A8	厨		
5704	穉	2	8D73	6C0A	穉		氈，異體字問題。
5706	敘	28	8D75	654D	敘		敘，異體字問題。
5742	澁	2	8DBB	6F81	澁		
5753	潜	3	8DC6	6F5C	潜		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5768	濶	9	8DD5	6FF6	濶		
5797	灑	1	8DF2	7054	灑		

## 二、手動取代表說明：

《繡像金瓶梅詞話》檔案中需手動修改的情況分為以下幾種：

(一) 《繡像金瓶梅詞話》檔案中之舊版造字，因字形部份不同而無法全部以程式取代，皆手動修改成Windows XP能顯示的字或構字式，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. 取代：實際使用字形與舊版造字不同時，於此欄列出。
4. 檔案原文摘錄：摘錄檔案《繡像金瓶梅詞話》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。
5. 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。
6. 備註：記錄修改相同辭句的次數或其他事項。

### 附件二、《繡像金瓶梅詞話》手動取代表

編號	造字	取代	檔案原文摘錄	手動取代結果	備註
2206	祆	祆	他則待__廟火	他則待 <u>祆</u> 廟火	
4237	髧	髧𠄎八	<u>羞</u> ，老婆無不曲	<u>髧𠄎八</u> ，老婆無不曲	
			<u>羞</u> 的肥一般	<u>髧𠄎八</u> 的肥一般	

(二)《繡像金瓶梅詞話》檔案中，對於舊版造字以外發現的錯字，以手動修改，並製作錯字修改表，請參考附件三。欄位說明如下：

1. 檔案原文摘錄：摘錄檔案《繡像金瓶梅詞話》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。
2. 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。
3. 備註：記錄修改相同詞句的次數或其他事項。

### 附件三、《繡像金瓶梅詞話》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
卓椅□□	卓椅 <u>錦杌</u>	
那件玄色□	那件玄色 <u>□</u>	
這□骨禿了	這 <u>狗</u> 骨禿了	
桂姐□□出來	桂姐 <u>也走</u> 出來	
冤家作怪□氣	冤家作怪 <u>搗</u> 氣	
頭 <u>下</u> 青天自恁欺	頭 <u>上</u> 青天自恁欺	
乞\023疆衫i鏡叟訴冤	乞 <u>臙肉磨</u> 鏡叟訴冤	
<u>捎</u> 著腳兒就	<u>息</u> 著腳兒就	