

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document
Title: Proposal to encode seven additional Tangut components
Source: Andrew West
Status: Individual Contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Date: 2018-06-01

1. Introduction

The Tangut Components block contains 755 characters for Tangut components that are used as radicals and/or IDS components in modern Tangut scholarship. When this set of components was proposed for encoding in September 2014 (see WG2 N4636 or L2/14-228) there were no known examples of Tangut components used in original Tangut sources. However, in November 2014 a woodblock edition of a previously unknown Tangut linguistic text entitled *Essential Selection of Often-Transmitted Homonyms and Mixed Characters* 𐰇𐰆𐰏𐰐𐰑𐰒𐰓𐰔𐰕𐰖𐰗𐰘𐰙𐰚𐰛𐰜𐰝𐰞𐰟𐰠𐰡𐰢𐰣𐰤𐰥𐰦𐰧𐰨𐰩𐰪𐰫𐰬𐰭𐰮𐰯𐰰𐰱𐰲𐰳𐰴𐰵𐰶𐰷𐰸𐰹𐰺𐰻𐰼𐰽𐰾𐰿 (Zéyào Chángchuán Tóngmíng Zázì 擇要常傳同名雜字 in Chinese translation, referred to in short as *Homonyms* below) was sold at auction in Beijing. This text includes two folios listing Tangut components ordered by position or type (see Figs. 1 and 2). In November 2017, thanks to the kindness of Prof. Nie Hongyin, I was able to examine photographs of the entire text, and map the components shown in this book to the set of encoded Tangut components (see Table 1). Seven Tangut components listed in this work are not currently encoded, and are required by scholars studying this important text.

2. Tangut Components

The newly-discovered *Homonyms* text includes 230 entries for individual Tangut components, of which entries 1 through 188 show components that occur at the top (nos. 1–31), left or right side (nos. 32–158), middle (nos. 159–169), and exclusively on the right side (nos. 170–188), each with a sample character in regular script and cursive script demonstrating the component. Two entries (nos. 55 and 74) show left side and right side forms of the component, so there are 190 components in total. At the end of the second folio there are also four components that extend underneath from the left (nos. 227–230), but with a description instead of a sample character. 189 or these 194 components are already encoded in the Tangut Components block, but five components (nos. 1, 71, 105, 112, and 140) are not encoded.

In addition to the individual components, the text lists 38 Tangut ideographs that also occur as lateral components in other ideographs (nos. 189–226). Of these 38 characters, 35 are encoded both as Tangut ideographs and as Tangut components, one is uncertain (no. 216), and two are encoded as Tangut ideographs but not as components (nos. 205 and 222), even though they both occur as right side components in other Tangut ideographs.

Fig. 1: Homonyms folio 4

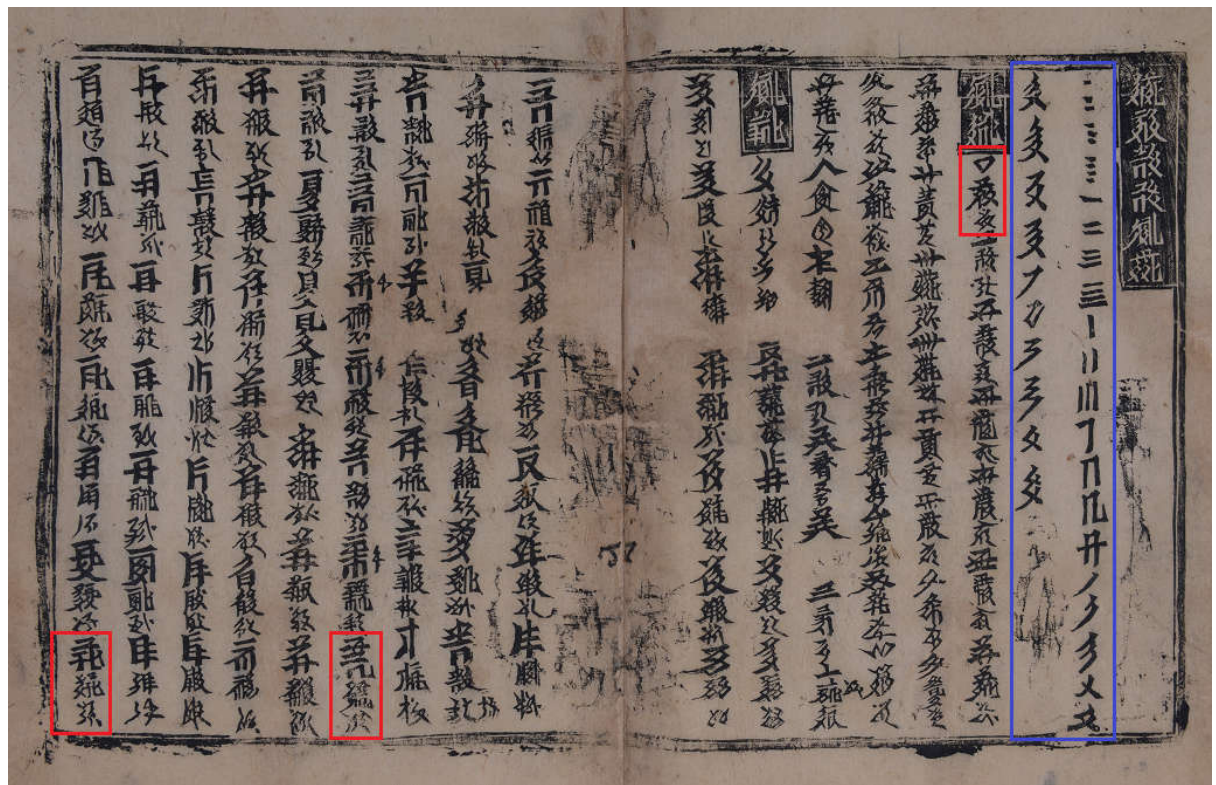


Fig. 2: Homonyms folio 5

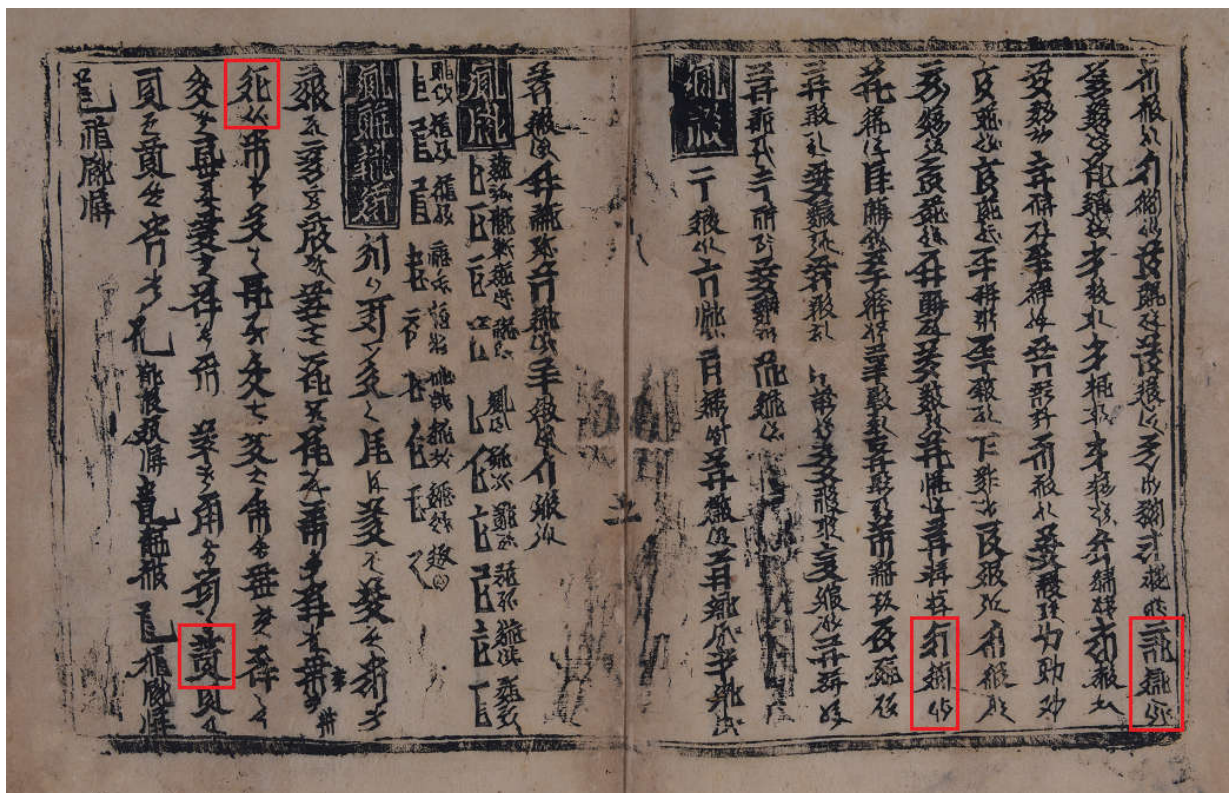


Table 1: Tangut Components in Homonyms

No.	Position	Component	C/P	Character Name	Sample	Cursive
1	04A401				𐰇	𐰇
2	04A402	𐰇	U+1881B	TANGUT COMPONENT-028	𐰇	𐰇
3	04A403	𐰇	U+188B8	TANGUT COMPONENT-185	𐰇	𐰇
4	04A404	𐰇	U+18944	TANGUT COMPONENT-325	𐰇	𐰇
5	04A405	𐰇	U+18949	TANGUT COMPONENT-330	𐰇	𐰇
6	04A406	𐰇	U+189C9	TANGUT COMPONENT-458	𐰇	𐰇
7	04A407	𐰇	U+189AF	TANGUT COMPONENT-432	𐰇	𐰇
8	04A501	𐰇	U+18A1D	TANGUT COMPONENT-542	𐰇	𐰇
9	04A502	𐰇	U+1882C	TANGUT COMPONENT-045	𐰇	𐰇
10	04A503	𐰇	U+18869	TANGUT COMPONENT-106	𐰇	𐰇
11	04A504	𐰇	U+188D8	TANGUT COMPONENT-217	𐰇	𐰇
12	04A505	𐰇	U+18863	TANGUT COMPONENT-100	𐰇	𐰇
13	04A506	𐰇	U+188CD	TANGUT COMPONENT-206	𐰇	𐰇
14	04A507	𐰇	U+18845	TANGUT COMPONENT-070	𐰇	𐰇
15	04A508	𐰇	U+1888E	TANGUT COMPONENT-143	𐰇	𐰇
16	04A601	𐰇	U+18907	TANGUT COMPONENT-264	𐰇	𐰇

No.	Position	Component	C/P	Character Name	Sample	Cursive
17	04A602	𠂇	U+188CB	TANGUT COMPONENT-204	𠂇	𠂇
18	04A603	𠂈	U+1885B	TANGUT COMPONENT-092	𠂈	𠂈
19	04A604	𠂉	U+18853	TANGUT COMPONENT-084	𠂉	𠂉
20	04A605	𠂊	U+188BB	TANGUT COMPONENT-188	𠂊	𠂊
21	04A606	𠂋	U+1885D	TANGUT COMPONENT-094	𠂋	𠂋
22	04A607	𠂌	U+189B4	TANGUT COMPONENT-437	𠂌	𠂌
23	04A608	𠂍	U+1884D	TANGUT COMPONENT-078	𠂍	𠂍
24	04A701	𠂎	U+1891B	TANGUT COMPONENT-284	𠂎	𠂎
25	04A702	𠂏	U+18817	TANGUT COMPONENT-024	𠂏	𠂏
26	04A703	𠂐	U+188BA	TANGUT COMPONENT-187	𠂐	
27	04A704	𠂑	U+188AB	TANGUT COMPONENT-172	𠂑	𠂑
28	04A705	𠂒	U+18934	TANGUT COMPONENT-309	𠂒	𠂒
29	04A706	𠂓	U+18A18	TANGUT COMPONENT-537		
30	04A707	𠂔	U+1884E	TANGUT COMPONENT-079	𠂔	𠂔
31	04A708	𠂕	U+18819	TANGUT COMPONENT-026	𠂕	𠂕
32	04A801	𠂖	U+18906	TANGUT COMPONENT-263	𠂖	𠂖
33	04A802	𠂗	U+18823	TANGUT COMPONENT-036	𠂗	

No.	Position	Component	C/P	Character Name	Sample	Cursive
34	04A803	𠂔	U+189E6	TANGUT COMPONENT-487	𠂔	𠂔
35	04A804	𠂕	U+18A57	TANGUT COMPONENT-600	𠂕	𠂕
36	04A805	𠂖	U+188AA	TANGUT COMPONENT-171	𠂖	𠂖
37	04A806	𠂗	U+18929	TANGUT COMPONENT-298	𠂗	𠂗
38	04A901	𠂘	U+1892C	TANGUT COMPONENT-301	𠂘	𠂘
39	04A902	𠂙	U+189B3	TANGUT COMPONENT-436	𠂙	𠂙
40	04A903	𠂚	U+189D3	TANGUT COMPONENT-468	𠂚	
41	04A904	𠂛	U+18A3C	TANGUT COMPONENT-573	𠂛	𠂛
42	04A905	𠂜	U+189A2	TANGUT COMPONENT-419	𠂜	𠂜
43	04A906	𠂝	U+18A0F	TANGUT COMPONENT-528	𠂝	𠂝
44	04A907	𠂞	U+18821	TANGUT COMPONENT-034	𠂞	𠂞
45	04B101	𠂟	U+188E9	TANGUT COMPONENT-234	𠂟	𠂟
46	04B102	𠂠	U+18862	TANGUT COMPONENT-099	𠂠	𠂠
47	04B103	𠂡	U+1897E	TANGUT COMPONENT-383	𠂡	𠂡
48	04B104	𠂢	U+1892E	TANGUT COMPONENT-303	𠂢	𠂢
49	04B105	𠂣	U+1886F	TANGUT COMPONENT-112	𠂣	𠂣
50	04B106	𠂤	U+189D8	TANGUT COMPONENT-473	𠂤	𠂤

No.	Position	Component	C/P	Character Name	Sample	Cursive
51	04B107	𠂔	U+188F2	TANGUT COMPONENT-243	𠂔	𠂔
52	04B201	𠂔	U+18903	TANGUT COMPONENT-260	𠂔	𠂔
53	04B202	𠂔	U+188C7	TANGUT COMPONENT-200	𠂔	𠂔
54	04B203	𠂔	U+18AA3	TANGUT COMPONENT-676	𠂔	𠂔
55	04B204	𠂔 𠂔	U+18A5A U+18A5B	TANGUT COMPONENT-603 TANGUT COMPONENT-604	𠂔	𠂔
56	04B205	𠂔	U+18AB4	TANGUT COMPONENT-693	𠂔	𠂔
57	04B206	𠂔	U+18A03	TANGUT COMPONENT-516	𠂔	𠂔
58	04B301	𠂔	U+1899E	TANGUT COMPONENT-415	𠂔	𠂔
59	04B302	𠂔	U+1886E	TANGUT COMPONENT-111	𠂔	𠂔
60	04B303	𠂔	U+18918	TANGUT COMPONENT-281	𠂔	
61	04B304	𠂔	U+18946	TANGUT COMPONENT-327	𠂔	𠂔
62	04B305	𠂔	U+18947	TANGUT COMPONENT-328	𠂔	𠂔
63	04B306	𠂔	U+18989	TANGUT COMPONENT-394	𠂔	𠂔
64	04B307	𠂔	U+1882D	TANGUT COMPONENT-046	𠂔	𠂔
65	04B401	𠂔	U+18A2F	TANGUT COMPONENT-560	𠂔	𠂔
66	04B402	𠂔	U+18A30	TANGUT COMPONENT-561	𠂔	𠂔
67	04B403	𠂔	U+18954	TANGUT COMPONENT-341	𠂔	𠂔

No.	Position	Component	C/P	Character Name	Sample	Cursive
68	04B404	𠂇	U+1896E	TANGUT COMPONENT-367	𠂇	𠂇
69	04B405	𠂈	U+189AE	TANGUT COMPONENT-431	𠂈	𠂈
70	04B406	𠂉	U+189EE	TANGUT COMPONENT-495	𠂉	𠂉
71	04B407	𠂊			𠂊	𠂊
72	04B501	𠂋	U+18939	TANGUT COMPONENT-314	𠂋	𠂋
73	04B502	𠂌	U+18AA2	TANGUT COMPONENT-675	𠂌	𠂌
74	04B503	𠂍 𠂎	U+18A58 U+18A59	TANGUT COMPONENT-601 TANGUT COMPONENT-602	𠂍	𠂍
75	04B504	𠂏	U+18A13	TANGUT COMPONENT-532	𠂏	𠂏
76	04B505	𠂐	U+18A20	TANGUT COMPONENT-545	𠂐	𠂐
77	04B506	𠂑	U+189B1	TANGUT COMPONENT-434	𠂑	𠂑
78	04B601	𠂒	U+18923	TANGUT COMPONENT-292	𠂒	𠂒
79	04B602	𠂓	U+189A0	TANGUT COMPONENT-417	𠂓	𠂓
80	04B603	𠂔	U+1899D	TANGUT COMPONENT-414	𠂔	𠂔
81	04B604	𠂕	U+18A08	TANGUT COMPONENT-521	𠂕	𠂕
82	04B605	𠂖	U+189A6	TANGUT COMPONENT-423	𠂖	𠂖
83	04B606	𠂗	U+1897C	TANGUT COMPONENT-381	𠂗	𠂗
84	04B607	𠂘	U+188EB	TANGUT COMPONENT-236	𠂘	𠂘

No.	Position	Component	C/P	Character Name	Sample	Cursive
85	04B701	𠂇	U+18955	TANGUT COMPONENT-342	𠂇	𠂇
86	04B702	𠂈	U+18A5F	TANGUT COMPONENT-608	𠂈	𠂈
87	04B703	𠂉	U+18836	TANGUT COMPONENT-055	𠂉	𠂉
88	04B704	𠂊	U+18878	TANGUT COMPONENT-121	𠂊	𠂊
89	04B705	𠂋	U+18875	TANGUT COMPONENT-118	𠂋	𠂋
90	04B706	𠂌	U+188F4	TANGUT COMPONENT-245	𠂌	𠂌
91	04B707	𠂍	U+18970	TANGUT COMPONENT-369	𠂍	𠂍
92	04B801	𠂎	U+18877	TANGUT COMPONENT-120	𠂎	𠂎
93	04B802	𠂏	U+188DC	TANGUT COMPONENT-221	𠂏	𠂏
94	04B803	𠂐	U+18964	TANGUT COMPONENT-357	𠂐	𠂐
95	04B804	𠂑	U+188DD	TANGUT COMPONENT-222	𠂑	𠂑
96	04B805	𠂒	U+188E0	TANGUT COMPONENT-225	𠂒	𠂒
97	04B806	𠂓	U+18A51	TANGUT COMPONENT-594	𠂓	𠂓
98	04B807	𠂔	U+188F9	TANGUT COMPONENT-250	𠂔	𠂔
99	04B901	𠂕	U+1894C	TANGUT COMPONENT-333	𠂕	𠂕
100	04B902	𠂖	U+188F8	TANGUT COMPONENT-249	𠂖	𠂖
101	04B903	𠂗	U+188DE	TANGUT COMPONENT-223	𠂗	𠂗

No.	Position	Component	C/P	Character Name	Sample	Cursive
102	04B904	𠂔	U+18966	TANGUT COMPONENT-359	𠂔	𠂔
103	04B905	𠂕	U+1894D	TANGUT COMPONENT-334	𠂕	𠂕
104	04B906	𠂖	U+189E3	TANGUT COMPONENT-484	𠂖	𠂖
105	04B907	𠂗			𠂗	𠂗
106	05A101	𠂘	U+18974	TANGUT COMPONENT-373	𠂘	𠂘
107	05A102	𠂙	U+18888	TANGUT COMPONENT-137	𠂙	𠂙
108	05A103	𠂚	U+18A06	TANGUT COMPONENT-519	𠂚	𠂚
109	05A104	𠂛	U+18A7C	TANGUT COMPONENT-637	𠂛	𠂛
110	05A105	𠂜	U+18820	TANGUT COMPONENT-033	𠂜	𠂜
111	05A106	𠂝	U+18831	TANGUT COMPONENT-050	𠂝	𠂝
112	05A107	𠂞			𠂞	𠂞
113	05A201	𠂟	U+18994	TANGUT COMPONENT-405	𠂟	𠂟
114	05A202	𠂠	U+18A10	TANGUT COMPONENT-529	𠂠	𠂠
115	05A203	𠂡	U+18855	TANGUT COMPONENT-086	𠂡	𠂡
116	05A204	𠂢	U+188A9	TANGUT COMPONENT-170	𠂢	𠂢
117	05A205	𠂣	U+1881E	TANGUT COMPONENT-031	𠂣	𠂣
118	05A206	𠂤	U+1890C	TANGUT COMPONENT-269	𠂤	𠂤

No.	Position	Component	C/P	Character Name	Sample	Cursive
119	05A207		U+188BF	TANGUT COMPONENT-192		
120	05A301		U+18A36	TANGUT COMPONENT-567		
121	05A302		U+18986	TANGUT COMPONENT-391		
122	05A303		U+18953	TANGUT COMPONENT-340		
123	05A304		U+18A34	TANGUT COMPONENT-565		
124	05A305		U+188C2	TANGUT COMPONENT-195		
125	05A306		U+18AB1	TANGUT COMPONENT-690		
126	05A307		U+1883D	TANGUT COMPONENT-062		
127	05A401		U+1898C	TANGUT COMPONENT-397		
128	05A402		U+189F9	TANGUT COMPONENT-506		
129	05A403		U+188B7	TANGUT COMPONENT-184		
130	05A404		U+18940	TANGUT COMPONENT-321		
131	05A405		U+18885	TANGUT COMPONENT-134		
132	05A406		U+188E3	TANGUT COMPONENT-228		
133	05A407		U+18902	TANGUT COMPONENT-259		
134	05A501		U+18872	TANGUT COMPONENT-115		
135	05A502		U+18951	TANGUT COMPONENT-338		

No.	Position	Component	C/P	Character Name	Sample	Cursive
136	05A503	𠂔	U+18941	TANGUT COMPONENT-322	𠂔	𠂔
137	05A504	𠂔	U+18A1A	TANGUT COMPONENT-539	𠂔	𠂔
138	05A505	𠂔	U+18997	TANGUT COMPONENT-408	𠂔	𠂔
139	05A506	𠂔	U+1896B	TANGUT COMPONENT-364	𠂔	𠂔
140	05A507	𠂔			𠂔	𠂔
141	05A601	𠂔	U+189B2	TANGUT COMPONENT-435	𠂔	𠂔
142	05A602	𠂔	U+18973	TANGUT COMPONENT-372	𠂔	𠂔
143	05A603	𠂔	U+189E8	TANGUT COMPONENT-489	𠂔	𠂔
144	05A604	𠂔	U+18A87	TANGUT COMPONENT-648	𠂔	𠂔
145	05A605	𠂔	U+189E5	TANGUT COMPONENT-486	𠂔	𠂔
146	05A606	𠂔	U+189FF	TANGUT COMPONENT-512	𠂔	𠂔
147	05A607	𠂔	U+1894A	TANGUT COMPONENT-331	𠂔	𠂔
148	05A701	𠂔	U+189C0	TANGUT COMPONENT-449	𠂔	𠂔
149	05A702	𠂔	U+18ABC	TANGUT COMPONENT-701	𠂔	𠂔
150	05A703	𠂔	U+18A92	TANGUT COMPONENT-659	𠂔	𠂔
151	05A704	𠂔	U+1898D	TANGUT COMPONENT-398	𠂔	𠂔
152	05A705	𠂔	U+18952	TANGUT COMPONENT-339	𠂔	𠂔

No.	Position	Component	C/P	Character Name	Sample	Cursive
153	05A706	𪛗	U+1898A	TANGUT COMPONENT-395	𪛗	𪛗
154	05A707	𪛘	U+18937	TANGUT COMPONENT-312	𪛘	𪛘
155	05A801	𪛙	U+18A26	TANGUT COMPONENT-551	𪛙	𪛙
156	05A802	𪛚	U+18899	TANGUT COMPONENT-154	𪛚	𪛚
157	05A803	𪛛	U+18A8C	TANGUT COMPONENT-653	𪛛	𪛛
158	05A804	𪛜	U+18A05	TANGUT COMPONENT-518	𪛜	𪛜
159	05A901	𪛝	U+18827	TANGUT COMPONENT-040	𪛝	𪛝
160	05A902	𪛞	U+1889A	TANGUT COMPONENT-155	𪛞	𪛞
161	05A903	𪛟	U+188DA	TANGUT COMPONENT-219	𪛟	𪛟
162	05A904	𪛠	U+18A19	TANGUT COMPONENT-538	𪛠	𪛠
163	05A905	𪛡	U+189BF	TANGUT COMPONENT-448	𪛡	𪛡
164	05A906	𪛢	U+188AF	TANGUT COMPONENT-176	𪛢	𪛢
165	05B101	𪛣	U+18A81	TANGUT COMPONENT-642	𪛣	𪛣
166	05B102	𪛤	U+188FF	TANGUT COMPONENT-256	𪛤	𪛤
167	05B103	𪛥	U+1891C	TANGUT COMPONENT-285	𪛥	𪛥
168	05B104	𪛦	U+18931	TANGUT COMPONENT-306	𪛦	𪛦
169	05B105	𪛧	U+18840	TANGUT COMPONENT-065	𪛧	𪛧

No.	Position	Component	C/P	Character Name	Sample	Cursive
170	05B201	乚	U+18822	TANGUT COMPONENT-035	𪛗	𪛗
171	05B202	𠂇	U+18834	TANGUT COMPONENT-053	𪛘	𪛘
172	05B203	𠂈	U+18867	TANGUT COMPONENT-104	𪛙	𪛙
173	05B204	𠂉	U+1899C	TANGUT COMPONENT-413	𪛚	𪛚
174	05B205	乚	U+18806	TANGUT COMPONENT-007	𪛛	𪛛
175	05B206	𠂇	U+18846	TANGUT COMPONENT-071	𪛜	𪛜
176	05B207	𠂈	U+1889E	TANGUT COMPONENT-159	𪛝	𪛝
177	05B208	𠂉	U+188B6	TANGUT COMPONENT-183	𪛞	𪛞
178	05B209	𠂊	U+18911	TANGUT COMPONENT-274	𪛟	𪛟
179	05B210	𠂋	U+188EE	TANGUT COMPONENT-239	𪛠	𪛠
180	05B301	𠂌	U+188C8	TANGUT COMPONENT-201	𪛡	𪛡
181	05B302	𠂍	U+189C8	TANGUT COMPONENT-457	𪛢	𪛢
182	05B303	𠂎	U+18A31	TANGUT COMPONENT-562	𪛣	𪛣
183	05B304	𠂏	U+1893E	TANGUT COMPONENT-319	𪛤	𪛤
184	05B305	𠂐	U+188ED	TANGUT COMPONENT-238	?	𪛦
185	05B306	𠂑	U+18874	TANGUT COMPONENT-117	𪛧	𪛧
186	05B307	𠂒	U+18895	TANGUT COMPONENT-150	𪛨	𪛨

No.	Position	Component	C/P	Character Name	Sample	Cursive
187	05B308	𠂇	U+188D2	TANGUT COMPONENT-211	𠂇	𠂇
188	05B309	𠂈	U+18809	TANGUT COMPONENT-010	𠂈	𠂈
189	05B401	𠂉 = 𠂉	U+17F03 U+18905	TANGUT IDEOGRAPH-17F03 TANGUT COMPONENT-262		𠂉
190	05B402	𠂊 = 𠂊	U+18229 U+18969	TANGUT IDEOGRAPH-18229 TANGUT COMPONENT-362		𠂊
191	05B403	𠂋 = 𠂋	U+178A8 U+1888C	TANGUT IDEOGRAPH-178A8 TANGUT COMPONENT-141		𠂋
192	05B404	𠂌 = 𠂌	U+17E6D U+188F4	TANGUT IDEOGRAPH-17E6D TANGUT COMPONENT-245		𠂌
193	05B405	𠂍 = 𠂍	U+186F5 U+18A84	TANGUT IDEOGRAPH-186F5 TANGUT COMPONENT-645		𠂍
194	05B406	𠂎 = 𠂎	U+18768 U+18AB2	TANGUT IDEOGRAPH-18768 TANGUT COMPONENT-691		𠂎
195	05B407	𠂏 = 𠂏	U+18086 U+18927	TANGUT IDEOGRAPH-18086 TANGUT COMPONENT-296		𠂏
196	05B501	𠂐 = 𠂐	U+187E0 U+18AE9	TANGUT IDEOGRAPH-187E0 TANGUT COMPONENT-746		𠂐
197	05B502	𠂑 = 𠂑	U+1821A U+1895F	TANGUT IDEOGRAPH-1821A TANGUT COMPONENT-352		𠂑
198	05B503	𠂒 = 𠂒	U+187C0 U+18AD6	TANGUT IDEOGRAPH-187C0 TANGUT COMPONENT-727		𠂒
199	05B504	𠂓 = 𠂓	U+1878C U+18ABF	TANGUT IDEOGRAPH-1878C TANGUT COMPONENT-704		𠂓
200	05B505	𠂔 = 𠂔	U+17007 U+18A55	TANGUT IDEOGRAPH-17007 TANGUT COMPONENT-598		𠂔
201	05B506	𠂕 = 𠂕	U+18350 U+189A6	TANGUT IDEOGRAPH-18350 TANGUT COMPONENT-423		𠂕
202	05B507	𠂖 = 𠂖	U+18527 U+189ED	TANGUT IDEOGRAPH-18527 TANGUT COMPONENT-494		𠂖
203	05B508	𠂗 = 𠂗	U+18570 U+18A0E	TANGUT IDEOGRAPH-18570 TANGUT COMPONENT-527		𠂗

No.	Position	Component	C/P	Character Name	Sample	Cursive
204	05B509	𠂇 = 𠂇	U+18790 U+18AC2	TANGUT IDEOGRAPH-18790 TANGUT COMPONENT-707		𠂇𠂇𠂇
205	05B601	𠂇 = 𠂇	U+178AC	TANGUT IDEOGRAPH-178AC		𠂇
206	05B602	𠂇 = 𠂇	U+182DD U+18998	TANGUT IDEOGRAPH-182DD TANGUT COMPONENT-409		𠂇
207	05B603	𠂇 = 𠂇	U+180AA U+18929	TANGUT IDEOGRAPH-180AA TANGUT COMPONENT-298		𠂇
208	05B604	𠂇 = 𠂇	U+17E15 U+188E0	TANGUT IDEOGRAPH-17E15 TANGUT COMPONENT-225		𠂇
209	05B605	𠂇 = 𠂇	U+17F86 U+1890A	TANGUT IDEOGRAPH-17F86 TANGUT COMPONENT-267		𠂇
210	05B606	𠂇 = 𠂇	U+18624 U+18A37	TANGUT IDEOGRAPH-18624 TANGUT COMPONENT-568		𠂇
211	05B607	𠂇 = 𠂇	U+1825F U+1897A	TANGUT IDEOGRAPH-1825F TANGUT COMPONENT-379		𠂇
212	05B608	𠂇 = 𠂇	U+187BC U+18AD5	TANGUT IDEOGRAPH-187BC TANGUT COMPONENT-726		𠂇
213	05B609	𠂇 = 𠂇	U+181DC U+1894F	TANGUT IDEOGRAPH-181DC TANGUT COMPONENT-336		𠂇
214	05B701	𠂇 = 𠂇	U+1828F U+1897F	TANGUT IDEOGRAPH-1828F TANGUT COMPONENT-384		𠂇
215	05B702	𠂇 = 𠂇	U+18797 U+18AC6	TANGUT IDEOGRAPH-18797 TANGUT COMPONENT-711		𠂇
216	05B703	𠂇 = 𠂇?	U+17C56 U+18ABE	TANGUT IDEOGRAPH-17C56 TANGUT COMPONENT-703		𠂇
217	05B704	𠂇 = 𠂇	U+185D0 U+18A1F	TANGUT IDEOGRAPH-185D0 TANGUT COMPONENT-544		𠂇
218	05B705	𠂇 = 𠂇	U+184C9 U+189D6	TANGUT IDEOGRAPH-184C9 TANGUT COMPONENT-471		
219	05B706	𠂇 = 𠂇	U+186CD U+18A63	TANGUT IDEOGRAPH-186CD TANGUT COMPONENT-612		𠂇
220	05B707	𠂇 = 𠂇	U+1866C U+18A3B	TANGUT IDEOGRAPH-1866C TANGUT COMPONENT-572		𠂇

No.	Position	Component	C/P	Character Name	Sample	Cursive
221	05B708	𠂇 = 𠂇	U+1835D U+189AB	TANGUT IDEOGRAPH-1835D TANGUT COMPONENT-428		𠂇
222	05B709	𠂈 = 𠂈	U+17368	TANGUT IDEOGRAPH-17368		
223	05B710	𠂉 = 𠂉	U+1824B U+18977	TANGUT IDEOGRAPH-1824B TANGUT COMPONENT-376		𠂉
224	05B801	𠂊 = 𠂊	U+184EC U+189E4	TANGUT IDEOGRAPH-184EC TANGUT COMPONENT-485		𠂊
225	05B802	𠂋 = 𠂋	U+187C7 U+18AD9	TANGUT IDEOGRAPH-187C7 TANGUT COMPONENT-730		𠂋
226	05B803	𠂌 = 𠂌	U+18302 U+1899E	TANGUT IDEOGRAPH-18302 TANGUT COMPONENT-415		𠂌
227	05B804	𠂍	U+188A6	TANGUT COMPONENT-167	𠂍𠂍𠂍𠂍	
228	05B805	𠂎	U+18AAB	TANGUT COMPONENT-684	𠂎𠂎	
229	05B806	𠂏	U+18A32	TANGUT COMPONENT-563	𠂏𠂏𠂍	
230	05B901	𠂐	U+18A0B	TANGUT COMPONENT-524	𠂐𠂏𠂍	

#205 𠂑 occurs as a right side components in 26 characters: U+17149 𠂑, U+1730C 𠂑, U+17341 𠂑, U+1759F 𠂑, U+17863 𠂑, U+17AAC 𠂑, U+17B2C 𠂑, U+17B33 𠂑, U+17B92 𠂑, U+17B95 𠂑, U+17C1E 𠂑, U+17CC7 𠂑, U+17F52 𠂑, U+17FF6 𠂑, U+18010 𠂑, U+180A1 𠂑, U+180F6 𠂑, U+1814F 𠂑, U+181E7 𠂑, U+1821D 𠂑, U+18282 𠂑, U+182F0 𠂑, U+1847A 𠂑, U+1851C 𠂑, U+186B6 𠂑, U+18746 𠂑.

#216 is not clear. It could be 𠂒 but this character does not occur as a component, or it could be the component 𠂒 but this component does not occur as a character in its own right.

#222 𠂓 occurs as a right side component in five characters: U+1766B 𠂓, U+17A48 𠂓, U+17EEF 𠂓, U+17FE2 𠂓, U+1843F 𠂓.

3. Thirty Letter Signs

The *Homonyms* text also lists thirty signs labelled as “thirty letters” 𪚩𪚪𪚫𪚬 (highlighted in blue in Fig. 1). These thirty signs can be divided into nine series of between two and four signs, with each series showing an incremental increase in stroke complexity (see Table 2).

Table 2: “30 letters” listed in the *Homonyms*

Series	Nos.	30 Letters
1	1-4	一 二 三 四
2	5-8	一 二 三 四
3	9-11	丨 丨 丨
4	12-15	𪚩 𪚪 𪚫 𪚬
5	16-18	𪚭 𪚮 𪚯
6	19-22	𪚰 𪚱 𪚲 𪚳
7	23-24	𪚴 𪚵
8	25-28	𪚶 𪚷 𪚸 𪚹
9	29-30	𪚺 𪚻

Series 1–3 consist of one through four short horizontal, long horizontal, and vertical strokes resembling tally marks, whereas Series 4–9 consist of a sequence of sub-components incrementally resulting in the components 𪚩 (U+18842), 𪚪 (U+18904), 𪚫 (U+188C0), 𪚬 (U+18805), and 𪚭 (U+1890A). The signs that correspond to encoded Tangut components are highlighted in yellow in Table 2. There are four sub-component signs that are not currently encoded as Tangut components (𪚮 𪚯 𪚰 𪚱).

Despite the title of this section, these thirty signs do not appear to be letters for a writing system, and as they do not occur anywhere else, their exact purpose is unknown. Therefore, none of the tally-like marks and unencoded sub-components are proposed for encoding at present.

4. Proposed Additions

The following seven Tangut components listed in the *Homonyms* are proposed for encoding. The proposed names simply extend the sequential numbers of the existing Tangut components.

Code Point	Glyph	Character Name
18AF3	𐰪	TANGUT COMPONENT-756
18AF4	𐰫	TANGUT COMPONENT-757
18AF5	𐰬	TANGUT COMPONENT-758
18AF6	𐰭	TANGUT COMPONENT-759
18AF7	𐰮	TANGUT COMPONENT-760
18AF8	𐰯	TANGUT COMPONENT-761
18AF9	𐰰	TANGUT COMPONENT-762

Unicode data properties:

18AF3; TANGUT COMPONENT-756; L₀; 0; L; ; ; ; ; N; ; ; ; ;

...

18AF9; TANGUT COMPONENT-762; L₀; 0; L; ; ; ; ; N; ; ; ; ;

All other Unicode character properties should be the same as for U+18800 TANGUT COMPONENT-001.

5. Bibliography

West, Andrew. "An introduction to the Tangut *Homonyms*". *Journal of Chinese Writing Systems*. June 2018 (Forthcoming). DOI: 10.1177/2513850218778825

6. Proposal Summary Form

ISO/IEC JTC 1/SC 2/WG 2 PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹. Please fill all the sections A, B and C below. Please read Principles and Procedures Document (P & P) from http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html for guidelines and details before filling this form. Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html. See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html for latest Roadmaps.
--

A. Administrative

1. Title:	Proposal to encode seven additional Tangut components		
2. Requester's name:	<i>Andrew West</i>		
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>		
4. Submission date:	<i>2018-06-01</i>		
5. Requester's reference (if applicable):			
6. Choose one of the following:			
This is a complete proposal:			<i>YES</i>
(or) More information will be provided later:			

B. Technical – General

1. Choose one of the following:			
a. This proposal is for a new script (set of characters):			<i>NO</i>
Proposed name of script:			
b. The proposal is for addition of character(s) to an existing block:			<i>YES</i>
Name of the existing block:	<i>Tangut Components</i>		
2. Number of characters in proposal:			<i>7</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):			
A-Contemporary	☐	B.1-Specialized (small collection)	☐
C-Major extinct	☒	D-Attested extinct	☐
F-Archaic Hieroglyphic or Ideographic	☐	E-Minor extinct	☐
		G-Obscure or questionable usage symbols	☐
4. Is a repertoire including character names provided?			
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?			<i>YES</i>
b. Are the character shapes attached in a legible form suitable for review?			<i>YES</i>
5. Fonts related:			
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?			
			<i>Andrew West</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):			
			<i>Andrew West</i>
6. References:			
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?			<i>YES</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?			<i>YES</i>
7. Special encoding issues:			
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?			<i>NO</i>
8. Additional Information:			
Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.			

¹ Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	NO
If YES explain	
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	YES
If YES, with whom?	
various Tangut experts	
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	NO
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	rare
Reference:	
5. Are the proposed characters in current use by the user community?	N/A
If YES, where? Reference:	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	NO
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	YES
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	NO
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	NO
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?	NO
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	NO
If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	NO
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	NO
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	