

IRGN1484a Annex S.1,2 (Version 4)Draft  
Annex S  
(informative)  
Procedure for the unification and arrangement  
of CJK Ideographs

The graphic character collections of CJK unified ideographs in ISO/IEC 10646 are specified in clause 33. They are derived from many more ideographs which are found in various different national and regional standards for coded character sets (the "sources").

This annex describes how the ideographs in this standard are derived from the sources by applying a set of unification procedures. It also describes how the ideographs in this standard are arranged in the sequence of consecutive code points to which they are assigned.

The source references for CJK unified ideographs are specified in clause 27.1.

Within the context of ISO/IEC 10646 a unification process is applied to the ideographic characters taken from the codes in the source groups. In this process, single ideographs from two or more of the source groups are associated together, and a single code point is assigned to them in this standard. The associations are made according to a set of procedures that are described below. Ideographs that are thus associated are described here as “unified”.

NOTE – The unification process does not apply to the following collections of ideographic characters:

CJK RADICALS SUPPLEMENT (2E80 - 2EFF)

KANGXI RADICALS (2F00 - 2FDF)

CJK COMPATIBILITY IDEOGRAPHS (F900 - FAFF with the exception of FA0E, FA0F, FA11, FA13, FA14, FA1F, FA21, FA23, FA24, FA27, FA28 and FA29)

CJK COMPATIBILITY IDEOGRAPHS SUPPLEMENT (2F800-

2FA1F).

## S.1 Unification procedure

### S.1.1 Scope of unification

Ideographs that are unrelated in historical derivation (non-cognate characters) have not been unified.

#### EXAMPLE

士, 土

NOTE – The difference of shape between the two ideographs in the above example is in the length of the lower horizontal line. This is considered an actual difference of shape. Furthermore these ideographs have different meanings. The meaning of the first is "Soldier" and of the second is "Soil or Earth".

An association between ideographs from different sources is made here if their shapes are sufficiently similar, according to the following system of classification.

### S.1.2 Two level classification

A two-level system of classification is used to differentiate (a) between abstract shapes and (b) between actual shapes determined by particular typefaces. Variant forms of an ideograph, which can not be unified, are identified based on the difference between their abstract shapes.

### S.1.3 Procedure

A unification procedure is used to determine whether two ideographs have the same abstract shape or different ones. The unification procedure has two stages, applied in the following order:

- a) Analysis of component structure;
- b) Analysis of component features;

### S.1.3.1 Analysis of component structure

In the first stage of the procedure the component structure of each ideograph is examined. A component of an ideograph is a geometrical combination of primitive elements. Alternative ideographs can be configured from the same set of components. Components can be combined to create a new component with a more complicated structure. An ideograph, therefore, can be defined as a component tree, where the top node is the ideograph itself, and the bottom nodes are the primitive elements. This is shown in Figure S.1.

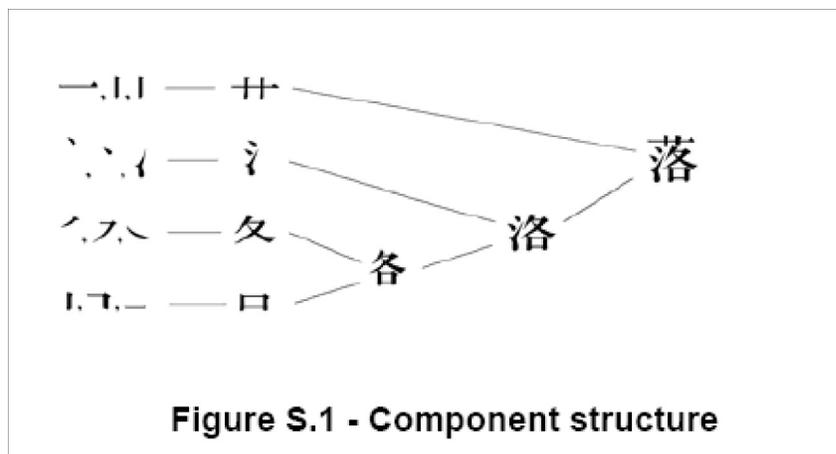


Figure S.1 - Component structure

Figure S.1 - Component structure

### S.1.3.2 Analysis of component

#### **features Comparison of component features**

In the second stage of the procedure, the components located at corresponding nodes of

two ideographs are compared, starting from the most superior node, as shown in Figure S.2.

The process of determining if two similar ideographs are unifiable begins with the top-most nodes of the component trees immediately below the root nodes. Considering the components at these nodes only, the ideographs are considered unifiable if and only if:

a) the number of components is the same,

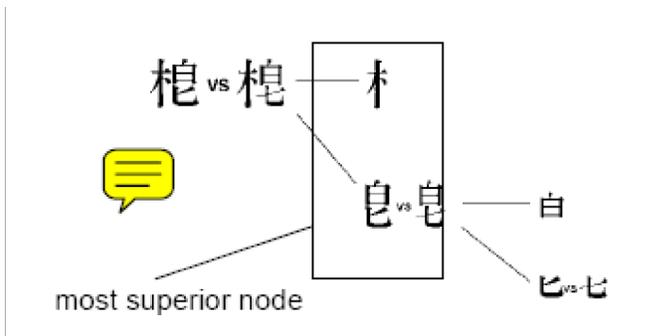
b) the relative positions of corresponding components within the full ideographs are the same, and

c) the corresponding components are themselves unifiable.

Note that some graphic elements used in drawing ideographs are unifiable when they occur as subcomponents of ideographs although they are not unifiable when they occur as independent ideographs.

If current unification rules are insufficient to determine whether or not the corresponding components are unifiable, the process may proceed recursively further along the component trees.

---



Remove and replace with the image below

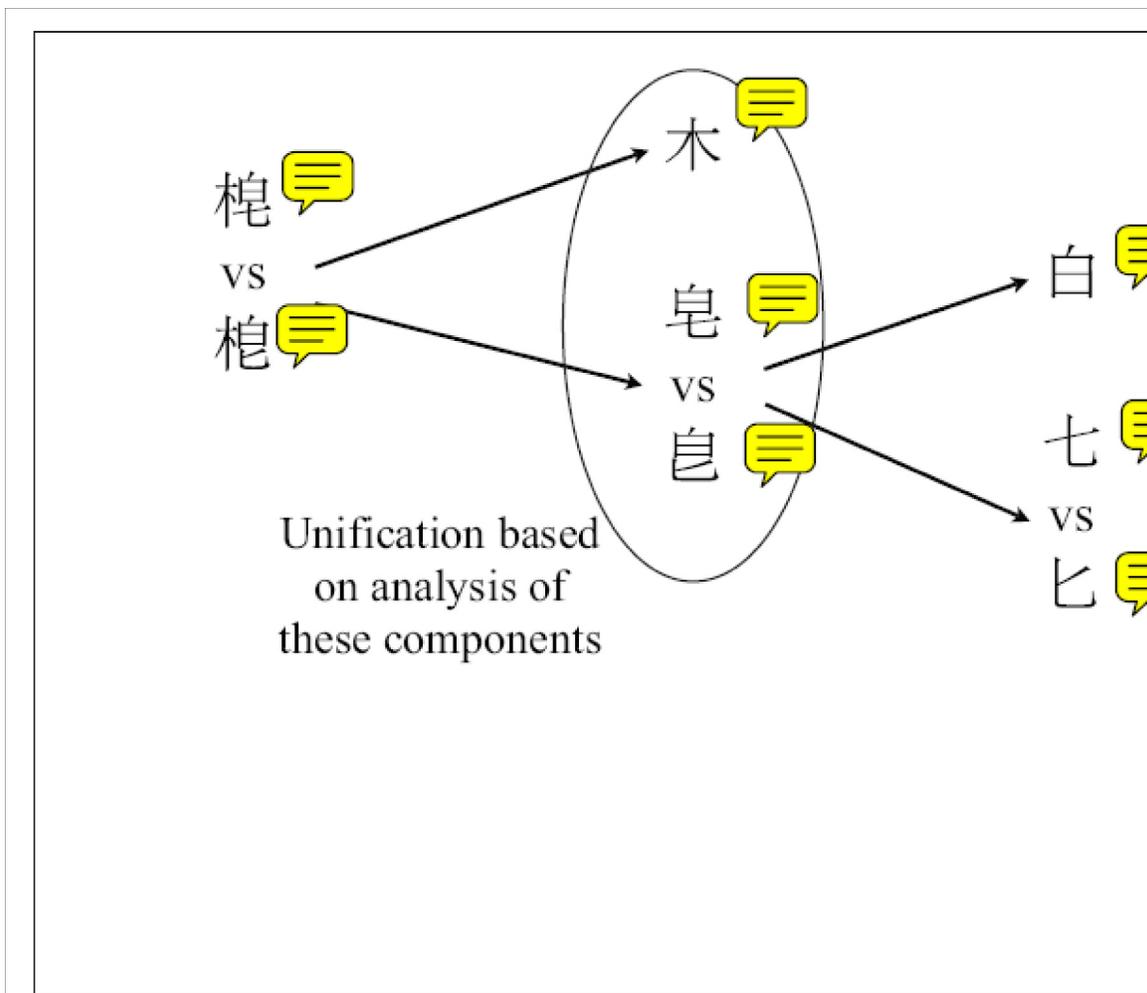


Figure S.2 - The most superior node of a component  
The following features of each ideograph to

be compared are examined:

- a) the number of components,
- b) the relative position of the components in each complete ideograph,
- c) the structure of corresponding components.

If one or more of the features a) to c) above are different between the ideographs in the comparison, the ideographs are considered to have different abstract shapes and are therefore not unified.

If all of the features a) to c) above are the same between the ideographs, the ideographs are considered to have the same abstract shape and are therefore unified.

Figure S.2 illustrates this comparison process. First, component trees are generated for each of the ideographs, 梶 and 梶. Next, the components at the highest nodes of the trees below the root nodes are considered. Each component tree has two nodes at this level, and the relative position of the components in the nodes is the same (one component at the left, and one at the right). Finally, the individual pairs of components are compared, in this case 木 vs 木 and 皂 vs 皂. The components in the former pair have the same abstract shape, but the components in the latter pair are not unified, so the two parent ideographs are not unified.

#### S.1.4 Examples of differences of abstract shapes

To illustrate rules derived from a) to c) in S.1.3.2, some typical examples of ideographs that are not unified, owing to differences of abstract shapes, are shown below.

##### S.1.4.1 Different number of components

The examples below illustrate rule a) since

the two ideographs in each pair have different numbers of components.

崖·厓, 肱·宏, 降·夆

S.1.4.2 Different relative positions of components

The examples below illustrate rule b).

Although the two ideographs in each pair have the same number of components, the relative positions of the components are different.

峰·峯, 荊·荆

When the left component can be made into an "L" shape, then "□XY" and "□XY" share the same abstract shape, as illustrated using the source glyphs for U+34F3:

052/243	𪗇	𪗇
<b>34F3</b>	3-3324	4-3052
	3-1904	4-1650

S.1.4.3 Different number of components Nonunifiable components

The examples below illustrate rule c). The structure of one (or more) corresponding components within the two ideographs in each pair is different, and they are considered nonunifiable.

扌·擴,	策·筭,	𠂇·𠂇,	又·𠂇,
僉·僉,	區·區,	夾·夾,	單·單,
隹·隹,	戔·戔,	贊·贊,	襄·襄,
隹·隹,	間·間,	朶·朶,	雋·雋,
恒·恆,	奂·奂,	人·人,	𠂇·𠂇,
𠂇·𠂇,			
又·又			

S.1.5 Differences of actual shapes  
 To illustrate the classification described in S.1.2, some typical examples of ideographs that are unified are shown below. The two or three ideographs in each group below have different actual shapes, but they are considered to have the same abstract shape, and are therefore unified.

迂·迂·迂,	示·示·示,	艮·艮·艮,	食·食·食,
黄·黄,	盥·盥,	曷·曷,	包·包,
青·青,	每·每,	册·册,	争·争,
畷·畷·畷,	录·录,	步·步,	者·者,
臭·臭,	并·并,	骨·骨,	吕·吕,
直·直,	鼎·鼎,	吴·吴·吴,	真·真·真,
爲·為,	单·单,	曾·曾·曾,	成·成,
專·專,	内·内,	晉·晋,	龜·龜,
++·++			

The differences are further classified according to the following examples.

a) Differences in rotated strokes/dots

半·半,	勺·勺,	羽·羽·羽,	酋·酋,
兼·兼,	益·益		

b) Differences in overshoot at the stroke initiation and/or termination

身·身,	雪·雪,	拐·拐,	不·不,
非·非,	周·周,	告·告	

c) Differences in contact of strokes

奧·奧,	酉·酉,	兕·兕,	查·查,
奔·奔			

d) Differences in protrusion at the folded corner of strokes

巨·巨

e) Differences in bent strokes

西·西

f) Differences in folding back at the stroke termination

朱·朱

g) Differences in accent at the stroke initiation

父·父, 丈·丈, 反·反

h) Differences in "rooftop" modification

八·八, 宀·宀

i) Combinations of the above differences

刃·刃·刃

j) Differences of a small part.

之·之·之, 步·步, 者·者  
專·專, 爲·為

These differences in actual shapes of a

unified ideograph are presented in the corresponding source columns for each code point entry in the code charts in clause 30 of this International Standard.

#### S.1.6 Source separation rule

To preserve data integrity through multiple stages of code conversion (commonly known as "round-trip integrity"), any ideographs that are separately encoded in any one of the source standards listed below have not been unified.

G-source: GB2312-80, GB12345-90, GB7589-87\*, GB7590-87\*, GB8565-88\*,

General Purpose Hanzi List for Modern Chinese Language\*

T-source: TCA-CNS 11643-1986/1st plane, TCA-CNS 11643-1986/2nd plane, TCA-CNS 11643-1986/14th plane\*

J-source: JIS X 0208-1990, JIS X 0212-1990

K-source: KS C 5601-1989, KS C 5657-1991

NOTE - A " \* " after the reference number of a standard indicates that some of the ideographs included in that standard are not introduced into the unified collection.

However, some ideographs encoded in two standards belonging to the same source group (e.g. GB2312-80 and GB12345-90) have been unified during the process of collecting ideographs from the source group.

The source separation rule described in this clause only applies to the CJK UNIFIED IDEOGRAPHS block

specified in the Basic Multilingual Plane.

NOTE - CJK Compatibility Ideographs are

created following a rule very similar to the source separation rule. However, the end result is the combination of a single CJK Unified Ideograph and one or several CJK Compatibility Ideographs. When the source separation rule is applied, all „similar“ source CJK Ideographs result in separate CJK Unified Ideographs.

## S.2 Arrangement procedure

### S.2.1 Scope of arrangement

The arrangement of the CJK UNIFIED IDEOGRAPHS in the code charts of clause 30 of this International Standard is based on the filing order of ideographs in the following dictionaries.

#### Priority Dictionary Edition

- 1 Kangxi Dictionary 康熙字典 Beijing 7th edition
- 2 Daikanwa Jiten 大漢和辭典 9th edition
- 3 Hanyu Dazidian 漢語大字典 1st edition
- 4 Daejaweon 大字源 1st edition

The dictionaries are used according to the priority order given in the table above. Priority 1 is highest. If an ideograph is found in one dictionary, the dictionaries of lower priority are not examined.

### S.2.2 Procedure

#### S.2.2.1 Ideographs found in the dictionaries

- a) If an ideograph is found in the Kangxi Dictionary, it is positioned in the code table in accordance with the Kangxi Dictionary order.
- b) If an ideograph is not found in the Kangxi Dictionary but is found in the Daikanwa Jiten, it is given a position at the end of

the radical-stroke group under which is indexed the nearest preceding Daikanwa Jiten character that also appears in the Kangxi dictionary.

c) If an ideograph is found in neither the Kangxi nor the Daikanwa, the Hanyu Dazidian and the Daejaweon dictionaries are referred to with a similar procedure.

S.2.2.2 Ideographs not found in the dictionaries

If an ideograph is not found in any of the four dictionaries, it is given a position at the end of the radicalstroke group (after the characters that are present in the dictionaries) and it is indexed under the same radical-stroke count.

IRGN1484b Annex S.3,4 (Version 4)  
 Draft  
 S.3 Source code separation examples  
 The pairs (or triplets) of ideographs shown below are exceptions to the unification rules described in S.1.  
 They are not unified because of the source separation rule described in S.1.6.

NOTE - The particular source group (or groups) that causes the source separation rule to apply is indicated by the letter (G, J, K, or T) that appears to the right of each pair (or triplet) of ideographs. The source groups that correspond to these letters are identified at the beginning of this annex.

丟	丟	T	兗	兗	T	单	单	T	国	国	T
4E1F	4E22		5156	5157		5355	5358		56EF	56FD	
么	么	GT	冊	冊	TJ	即	卽	TK	圈	圈	TJ
4E48	5E7A		518A	518C		5373	537D		5708	570F	
争	爭	GTJ	净	淨	G	卷	卷	TJ	圓	圓	T
4E89	722D		51C0	51C8		5377	5DFB		570E	5713	
仞	仞	J	尢	尢	T	叁	參	GT	圖	圖	T
4EDE	4EED		51E2	51E3		53C1	53C2		5716	5717	
併	併	T	刃	刃	TJ	參	參	T	垚	垚	T
4F75	5002		5203	5204		53C3	53C4		5759	5DE0	
侶	侶	T	刊	刊	TJ	吕	呂	T	埤	埤	J
4FA3	4FB6		520A	520B		5415	5442		57D2	57D3	
俣	俣	TJK	刪	刪	T	吞	吞	T	壑	壑	T
4FC1	4FE3		5220	522A		541E	5451		5848	588D	
俞	俞	T	別	別	T	吳	吳	TJ	填	填	TJ
4FDE	516A		5225	522B		5433	5434	5449	5861	586B	
俱	俱	T	券	券	TJ	呐	呐	T	增	增	T
4FF1	5036		5238	52B5		5436	5450		5897	589E	
值	值	T	刹	刹	T	告	告	T	壯	壯	GTJ
5024	503C		5239	524E		543F	544A		58EE	58EF	
偷	偷	T	勑	勑	T	啣	啣	T	壽	壽	T
5077	5078		524F	5259		5527	559E		58FD	5900	
偽	偽	TJ	剝	剝	T	喻	喻	T	夂	夂	T
507D	50DE		525D	5265		55A9	55BB		5910	657B	
兌	兌	T	劒	劒	J	嘘	嘘	T	夂	本	GTJ
514C	5151		5292	5294		5618	5653		5932	672C	
兔	兔	TJ	勻	勻	T	嚏	嚏	GTJ	奧	奧	J
514E	5154		52FB	5300		568F	5694		5965	5967	

獎	獎	獎
5968	596C	734E
妝	妝	
5986	599D	
妍	妍	
598D	59F8	
姍	姍	
59CD	59D7	
姬	姬	
59EB	59EC	
娛	娛	娛
5A1B	5A2F	5A31
婕	婕	
5A55	5AAB	
媪	媪	
5A7E	5AAE	
媪	媪	
5AAA	5ABC	
媪	媪	
5AAF	5B00	
媪	媪	
5B0E	5B14	
媪	媪	
5B24	5B37	
孳	孳	
5B73	5B76	
宮	宮	
5BAB	5BAE	
寬	寬	
5BDB	5BEC	
寧	寧	
5BDC	5BE7	
寢	寢	

TJ	專	專
	5C02	5C08
GT	將	將
	5C06	5C07
T	尔	尔
	5C13	5C14
T	尙	尙
	5C19	5C1A
GT	尫	尫
	5C2A	5C2B
T	檻	檻
	5C36	5C37
T	屏	屏
	5C4F	5C5B
T	崢	崢
	5CE5	5D22
TK	巔	巔
	5DD3	5DD4
T	併	併
	5E21	5E32
T	帶	帶
	5E2F	5E36
GT	并	并
	5E76	5E77
T	廐	廐
	5EC4	5ECF
T	弑	弑
	5F11	5F12
T	強	強
	5F37	5F3A
T	彈	彈
	5F39	5F3E
GTJ		

J	彡	彡
	5F50	5F51
GTJ	彡	彡
	5F54	5F55
T	彡	彡
	5F59	5F5A
T	彡	彡
	5F5B	5F5C
T	彡	彡
	5F5D	5F5E
T	彡	彡
	5F65	5F66
T	德	德
	5FB3	5FB7
GT	徵	徵
	5FB4	5FB5
T	惠	惠
	6075	60E0
T	悅	悅
	6085	60A6
TJ	悞	悞
	609E	60AE
T	息	息
	60B3	60EA
T	愠	愠
	6120	614D
T	慎	慎
	613C	614E
T	戩	戩
	6229	622C
T	戲	戲
	622F	6231
	戶	戶
		戶

TJ	戾	戾	6238
	623B	623E	
T	拋	拋	
	629B	62CB	
T	拔	拔	
	629C	62D4	
J	掄	掄	
	6329	635D	
T	插	插	插
	633F	63D2	63F7
T	捏	捏	
	634F	63D1	
T	搜	搜	
	635C	641C	
T	揭	揭	
	63B2	63ED	
TJ	搖	搖	搖
	63FA	6416	6447
T	搵	搵	
	63FE	6435	
T	擊	擊	
	6483	64CA	
T	教	教	
	654E	6559	
T	斂	斂	
	6553	655A	
TJ	既	既	
	65E2	65E3	
GT	昂	昂	
	6602	663B	
T	晚	晚	
	665A	6669	
T			

暨 暨  
 66A8 66C1  
 曾 曾  
 66FD 66FE  
 柺 柺  
 67B4 67FA  
 查 查  
 67E5 67FB  
 柵 柵  
 67F5 6805  
 稅 稅  
 68B2 68C1  
 榆 榆  
 6961 6986  
 概 概  
 6982 69EA  
 榼 榼  
 6985 69B2  
 榾 榾  
 699D 6A27  
 槓 槓  
 69C7 69D9  
 樣 樣  
 69D8 6A23  
 橫 橫  
 6A2A 6A6B  
 步 步  
 6B65 6B69  
 歲 歲  
 6B72 6B73  
 歿 歿  
 6B7F 6B81  
 殼 殼

6BBB 6BBC  
 T 毀 毀  
 6BC0 6BC1  
 J 每 每  
 6BCE 6BCF  
 T 氳 氳  
 6C32 6C33  
 T 污 污  
 6C5A 6C61  
 T 沒 沒  
 6C92 6CA1  
 T 淨 淨  
 6D44 6DE8  
 T 涉 涉  
 6D89 6E09  
 T 浼 浼  
 6D97 6D9A  
 T 淚 淚  
 6D99 6DDA  
 T 淥 淥  
 6DE5 6E0C  
 J 清 清  
 6DF8 6E05  
 TJ 渴 渴  
 6E07 6E34  
 T 溫 溫  
 6E29 6EAB  
 T 漚 漚  
 6E88 6F59  
 T 漑 漑  
 6E89 6F11  
 T 滾 滾  
 6EDA 6EFE  
 GTJ

潛 潛  
 6F5B 6FF3  
 T 瀨 瀨  
 7028 702C  
 T 為 為  
 70BA 7232  
 T 煢 煢  
 712D 7162  
 T 熙 熙  
 7155 7199  
 TJ 煨 煨  
 7174 7185  
 TJ 狀 狀  
 72B6 72C0  
 T 瑤 瑤  
 7464 7476  
 T 瓶 瓶  
 74F6 7501  
 T 產 產  
 7522 7523  
 T 瘦 瘦  
 75E9 7626  
 T 皤 皤  
 76A1 76A5  
 T 眞 眞  
 771E 771F  
 T 眾 眾  
 773E 8846  
 T 研 研  
 7814 784F  
 T 祿 祿  
 797F 7984  
 T 禿 禿

79BF 79C3  
 GTJK 稅 稅  
 7A05 7A0E  
 T 穗 穗  
 7A42 7A57  
 GTJ 箏 箏  
 7B5D 7B8F  
 GTJK 箏 箏  
 7BB3 7C08  
 J 篡 篡  
 7BE1 7C12  
 T 粵 粵  
 7CA4 7CB5  
 GT 絕 絕  
 7D55 7D76  
 TJ 綠 綠  
 7DA0 7DD1  
 T 緒 緒  
 7DD2 7DD6  
 T 緣 緣  
 7DE3 7E01  
 J 緼 緼  
 7DFC 7E15  
 T 緼 緼  
 7E48 7E66  
 TJ 羹 羹  
 7FAE 7FB9  
 TJ 翹 翹  
 7FF6 7FFA  
 T 胼 胼  
 80FC 8141  
 TJ 脫 脫  
 812B 8131  
 T

脍	脍	T	装	装	TJK	醞	醞	T	餅	餅	TJ
817D	8183		88C5	88DD		9196	919E		9905	9920	
烏	烏	GT	誾	誾	T	醬	醬	J	馱	馱	TJK
8203	8204		8A2E	8A7D		91A4	91AC		99B1	99C4	
舍	舍	TJ	說	說	T	鉞	鉞	T	駢	駢	TK
820D	820E		8AAA	8AAC		9203	9292		99E2	9A08	
舖	舖	J	諫	諫	TJ	銳	銳	T	飩	飩	T
8216	8217		8ACC	8AEB		92B3	92ED		9AA9	9AAB	
莊	莊	TJ	謠	謠	J	錄	錄	T	高	高	T
8358	838A		8B20	8B21		9304	9332		9AD8	9AD9	
菑	菑	TJ	豨	豨	T	鍊	鍊	TK	髮	髮	TJ
83D1	8458		8C5C	8C63		932C	934A		9AEA	9AEE	
蓋	蓋	T	走	走	TJ	鎮	鎮	TJ	鬪	鬪	T
8480	8495		8D70	8D71		93AD	93AE		9B2C	9B2D	
蔣	蔣	GJ	駢	駢	T	閱	閱	T	鯁	鯁	TJ
848B	8523		8EFF	8F27		95B1	95B2		9C1B	9C2E	
蔦	蔦	T	輜	輜	J	隍	隍	G	鳳	鳳	J
848D	853F		8F1C	8F3A		9667	9689		9CEF	9CF3	
蒞	蒞	T	輻	輻	T	青	青	T	鷓	鷓	J
8570	8580		8F3C	8F40		9751	9752		9D87	9DAB	
薰	薰	T	达	达	T	静	靜	GTJ	鷓	鷓	J
85AB	85B0		8FBE	8FD6		9759	975C		9DC6	9DCF	
蘊	蘊	T	迸	迸	TJ	靱	靱	J	麪	麪	T
85F4	860A		8FF8	902C		976D	9771		9EAA	9EAB	
虚	虚	T	遙	遙	J	頹	頹	T	麼	麼	T
865A	865B		9059	9065		9839	983D		9EBC	9EBD	
蛻	蛻	T	邢	邢	T	顏	顏	TJ	黄	黄	T
86FB	8715		90A2	90C9		984F	9854		9EC3	9EC4	
衛	衛	TJK	郎	郎	T	顛	顛	J	黑	黑	T
885B	885E		90CE	90DE		985A	985B		9ED1	9ED2	
袞	袞	TK	鄉	鄉	T	飲	飲	J			
886E	889E		90F7	9109	9115	98EE	98F2				

#### S.4 Nonunification examples

In accordance with the unification procedures described in S.1 the pairs (or triplets) of ideographs shown below are not unified. The reason for non-unification is indicated by the reference which appears to the right of each pair (or triplet). For “non -cognate” see S.1.1.  
NOTE - The reason for non-unification in these examples is different from the source separation rule described in clause S.1.6.

胃 胃 5191 80C4	non cognate	寶 寶 5BF3 5BF6	S.1.4.3	胸 胸 6710 80CA	non cognate	稻 稻 7A32 7A3B	S.1.4.3
冲 冲 51B2 6C96	S.1.4.3	廳 廳 5EF0 5EF3	S.1.4.1	眇 眇 6713 8101	Non cognate	翱 翱 7FF1 7FF6	S.1.4.3
決 決 51B3 6C7A	S.1.4.3	懷 懷 61D0 61F7	S.1.4.1	脞 脞 6718 8127	Non cognate	耆 耆 耆 8007 8008 8009	S.1.4.3
況 況 51B5 6CC1	S.1.4.3	斂 斂 6560 656A	S.1.4.3	腫 腫 6723 81A7	Non cognate	聽 聽 聽 8074 807C 807D	S.1.4.1
塚 塚 579B 579C	S.1.4.3	盼 盼 670C 80A6	non cognate	朵 朵 6735 6736	S.1.4.3	荊 荊 8346 834A	S.1.4.2
孳 孳 5B7C 5B7D	S.1.4.2	肫 肫 670F 80D0	non cognate	灑 灑 7054 7067	S.1.4.3	躲 躲 8EB1 8EB2	S.1.4.3

### S.5 Clarification of unification procedure

The examples provided in this annex **are not to be considered**

**exhaustive**, and therefore it is necessary to be able to apply the rules

and principles set forth in this annex to new situations. The rules,

principles, and illustrations that are provided in this annex are merely

descriptions of the calligraphic traditions and conventions used for

CJK ideographs, not a set of mathematical transformations.

When applying the unification rules and principles, the following

points should be considered:

- Non-cognate characters are not unified.
- Cognate characters, meaning that they share the same readings

and meanings, are unified according to the table below:

	Actual Shape	
	Different	Exact Match
Same Abstract Shape	Unify	Unify
Different Abstract Shape	Do Not Unify	

### S.5.1 Clarification of differences of actual shape

In accordance with the unification based on the analysis of components model of S.1.3 a pair of glyphs with different actual shape but the same abstract shape can be used as components to generate other pairs of glyphs that are generated in the same way. For example, given 兌•兑 have the same abstract shape then the following are also pairs with the same abstract shape:-

悅·悦, 挽·挽, 斂·斂, 稅·税, 浼·浼, 稅·税, 脫  
and 閱·閱.

Although (皞•皞) have the same abstract shape, and their corresponding components (皐•皐) also have the same abstract shape (cp. 翱•翱), this does not imply that the corresponding sub-components (白•白) also have the same abstract shape \*outside of this particular context\*. A relatively complex character is sometimes rather simplified (or otherwise modified from its normal independent form) when written small for use as a component in another character.

The following groups of ideographs shown below are further examples of differences of actual shape but with the same abstract shape when used as components:-

(Below added agreed new unification examples in the format shown)

耒 · 耒, 弱 · 弱, 𠂇 · 𠂇, 害 · 害, 勺 · 勺  
与 · 与, 唐 · 唐, 冉 · 冉, 寧 · 寧, 囟 · 囟  
鬲 · 鬲, 灰 · 灰, 華 · 華, 叟 · 叟 · 叟, 卑  
瓜 · 瓜, 𠂇 · 𠂇, 艮 · 艮 · 艮, 主 · 主, 敖  
止 · 止, 惠 · 惠, 夬 · 夬, 橐 · 橐, 壳 · 壳  
走 · 走, 尨 · 尨 · 尨, 取 · 取, 梟 · 梟, 虍  
畷 · 畷 · 畷, 巢 · 巢, 斗 · 斗, 微 · 微, 产  
會 · 會 · 會.

### S.5.2 Clarification of differences of abstract shape

In accordance with the model of S.1.3 different abstract shape with more than one used as components to generate at least one different abstract shape. For example, two different abstract shapes then it follows a different abstract shape.

Adding like components to pairs with different abstract shapes usually, but not always, results in pairs with the same abstract shape. Though as S.1.4.3 'Different structural components' illustrates, usually adding a pair of glyphs with different abstract shapes to a pair of different abstract shapes, therefore forming a pair of different abstract shapes then the resulting pairs, 嫺·嫺, 憫·憫, etc. However because the precedence in the S.1.3 model then it is not always the case that a pair of glyphs formed have the same abstract shape. For example though 口·厶 do not have the same abstract shape, adding 月 to both gives 冂·冂 which do have the same abstract shape.

The following groups of ideographs shown below are further examples of differences of abstract shape:-

(Below added agreed new non-unification examples in the format shown)

産•产, 尋•寻

In general, the Simplified components do not have the same abstract shape as their traditional counterparts. However the two ideographs 將 (5C07) and 将 (5C06) are unified when used as components.

## **S.6 Stability of unification rules**

With the consideration of additional encoded CJK ideographs, there is an on going need to consider new cases. Apart from the source separation rule, the rules and principles in this annex are stable, that is they apply to all CJK ideographs and do not change with time, and all new examples and sections must conform to these rules, and not contradict existing examples and sections.