

Considerations concerning the Chinese Root LGR

Last updated: October 28, 2016

1 Summary

This document analyzes the content of the current Chinese LGR as specified by the Chinese Generation Panel in terms of repertoire and variant sets. The latest draft of that Chinese LGR (CLGR7) is represented by the following files:

- CGP Proposal Draft 20161012.docx, [Proposal]
- Appendix I CGP LGR 20160923-mb.xml, [CLGR7]
- Appendix I CGP LGR 20160923-mb.html, (converted xml)

In evaluating this proposal, this document compares it to the dotAsia ZH set (see 2.4) which, like the CLGR7, attempts to cover both simplified and traditional Chinese labels.

While previous versions of these considerations analyzed in detail the content of the repertoire, the current status does not warrant such detailed analysis. The repertoire has matured and is now stable. However, there is still a concern whether the repertoire contains characters not needed in modern use and therefore there may be scope for making it smaller.

The following items summarize the analysis:

- The repertoire is now made of 19,746 code points, very close to the size of MSR-2 Hanzi set (19,850). It is now a full superset of the dotAsia set, itself made of 19,683 code points.
- It includes 2 characters not part of MSR-2, which will require a revision of MSR to include them.
- It is still not self-evident that a Chinese root LGR needs so many characters as are contained in the current draft. Many Chinese experts may consider a repertoire of 5,000 Han ideographs to reflect the repertoire recognized by a typical user. Even using the union of partially overlapping requirements for different communities, the total number of Han ideographs sufficient to express most of modern Chinese for purposes of root zone identifiers may well be smaller than that what is proposed here.
- **Accordingly, the IP is concerned about the current size of the CJK repertoire. The IP would like to remind the CJK community that the modern usage of each code point proposed in the LGR repertoire should be documented.**
- Notable progress was made in the variant description by systematically adding all 'reflexive-identity' mappings. These made the comparison with the dotAsia LGR much easier.

- Concerning variant sets, 111 out of over 3 500 variant sets containing 2 to 8 members show divergence between the CLGR and the comparison set (dotAsia ZH). This is less than 4% of the total amount.
- 60 out of the 111 variant sets that are different are the result of repertoire addition and are therefore expected. It is important to ensure that these additions follow the same principles that are used for the 3500 sets in general.
- Many of these 60 variant sets correspond to the addition of characters not essential for Chinese and it can be argued that these characters should be classified as out-of-repertoire and only use 'blocked' type variants in their mapping.
- Many of the remaining diverging 51 variant sets are still problematic because they show differences between CLGR7 and dotAsia, which is an established large Chinese domain registry that covers 99.7% of the same repertoire. CLGR's suggestion that the dotAsia variant sets are 'experimental' is very concerning.
- All parameters in this environment are large. The repertoire runs to over 19 000 code points, and more than 3 500 variant sets. Although the differences between the two sets are small in percentage, the number of differences is substantial in absolute terms.
- Most of the differences in variant sets found in a previous version of this feedback are still present and no explanation has yet been provided to explain in detail these differences.
- All these differences must be explored and deviations from existing practice should be justified. It is probably not possible to reach consensus on which differences to retain without some direct interaction with all concerned parties, including some face to face meetings.
- While the origin of unmodified variants can be traced to the dotAsia, dotcn, and dottw sets, no references or source information are provided for the modified or new variant sets. (The source for the unmodified variants, even if the IP could discover them, should be unambiguously documented as well – perhaps not on a per-variant level, but globally, with any exceptions prominently marked).
- Because this set must be integrated with the rest of the CJK sets (especially the Japanese set), it is important to get a version of this variant set which is agreeable to all concerned parties as soon as possible. From an integration perspective, the membership of the sets is paramount.
- Some of the additions concern ideographs that appear to be specific to Japanese or Korean. Their introduction to the CLGR7 necessarily creates additional variant mappings that will then affect the LGRs for Japanese and Korean. Again, maybe those should be only added as 'out-of-repertoire' type.
- These same additions (and additional variant mappings) would prevent the registration of labels under the Japanese or Korean LGRs that Chinese users might perceive as variants to Chinese labels, even if these additions might never be used in any labels applied for under the Chinese LGR. They would function effectively like out-of-repertoire variants, without being declared as such.
- Finally, there is a concern that creating too many 'trad' variants in a given variant set will overproduce allocatable labels. Reducing or eliminating these multiple 'trad' variants should be explored.

Outstanding Integration Panel recommendations:

The following recommendations that were made in a previous version of these considerations do not seem to be addressed in CLGR7:

- Provide documentation for the origin of the proposed variant mappings in CLGR7, particularly where they differ from established second level practice.
- Review variant sets that differ from second level practice and provide rationale for any differences.
- Review variant sets with multiple “trad” mappings to see if any of those could be changed to “blocked” to reduce the overproduction of allocatable variants.
- Document the specific requirements behind any decision to retain multiple “trad” variants.
- Provide detailed rationale for inclusion of J-specific or K-specific code points in this C-specific LGR. Please address the ramifications on variant sets deriving from these additions.

New Integration Panel recommendations:

The following recommendations are new to this document:

- Consider reducing the size of the LGR, using the principle that **unless modern** usage that is required in the context of internet identifiers can be documented for each code point proposed in the LGR repertoire, it should not be included. (Per RFC 6912 it is OK for the Root Zone to be more restrictive in this regard than other zones).
- If justification based on such usage has been established for inclusion, document such usage.
- When presenting special cases and deviations for variant sets in the LGR proposal document, consider presenting these variant sets in term of sets, not as separate code point based entries, to ensure that the sets are fully transitive and reflexive. Note that the XML LGR file is the reference for the full definition of these variant sets.
- If J-specific and K-specific code points are not included as full members of this repertoire, they should be included as out-of-repertoire variants and variant sets created accordingly (see main text for details).
- Provide references to all variant mappings using available sources such as Unihan, dotAsia, and any other relevant sources, using the ‘ref’ attribute on the ‘var’ element.

2 Definitions

2.1 CLGR7

The term (CLGR7) represents the Proposed Chinese root LGR under review here, both in terms of repertoire and variant sets defined in the XML file. The term CLGR6 may be used to represent the earlier version of the Chinese LGR.

2.2 IICORE collection

The International Ideographs Core (IICORE) is a fixed collection of CJK Ideographic code points deemed essential to all IRG Asian constituencies except Vietnam (a total of 7 sources). It contains 9 810 code points and is part of both ISO/IEC 10646 and Unicode. It was created by IRG based on priority (A to C, A being the highest) among its 7 sources.

2.3 MSR-2 CJK repertoire

The CJK repertoire in MSR-2 consists of 19 850 CJK Unified Ideographs, corresponding to the union of the following sub-repertoires:

- 1) dotAsia Japanese https://www.iana.org/domains/idn-tables/tables/asia_ja_1.1.txt
- 2) dotAsia Chinese https://www.iana.org/domains/idn-tables/tables/asia_zh_1.1.txt
- 3) IICORE as defined in Unicode 6.3
- 4) Code point U+9DC0.

The dotAsia Chinese repertoire is itself a union of repertoires from various Chinese sources such as China PRC, Hong Kong SARs, and Taiwan.

Note that MSR-2 also contains a few code points that have the ‘Han’ extended script property but are not considered CJK Ideographs (for example U+3005 IDEOGRAPHIC ITERATION MARK and U+3006 IDEOGRAPHIC CLOSING MARK).

2.4 dotAsia LGR

A transcription of the dotAsia (ZH) domain name definition available at https://www.iana.org/domains/idn-tables/tables/asia_zh_1.1.txt into the XML-format is publicly available for review at <https://www.icann.org/sites/default/files/packages/lgr/lgr-second-level-chinese-15may16-en.xml>. This transcription was created as part of reference for 2nd level domain. It shares many features with the proposed root Chinese LGR. The dotAsia table (or its XML transcription) represents an important set that can be used as a reference point for the comparison in terms of both the repertoire and the variant sets. It contains 19 684 Han ideographs and 3 505 variant sets. In comparison, the current Chinese Root Zone LGR draft (CLGR7) contains 19 746 Han ideographs and 3 518 variant sets.

Technically, the transcription of the dotAsia (ZH) domain contains one more CJK ideograph than dotAsian original table: U+9DC0 that was added to complete a variant set. The original IDN table contains 19 683 Han ideographs. In following comparisons of CLGR7 and dotAsia repertoire, the original 19 683 code points defined in dotAsia should be used.

The dotAsia repertoire is fully included in CLGR7. CLGR7 contains an additional 63 code points not included in dotAsia.

2.5 Unihan

The Unihan database at <http://www.unicode.org/charts/unihan.html> is a Unicode Standard component containing information related to all CJK Ideographs. That information includes sources, variants, dictionaries, etc. As such it is an extremely useful tool to validate the CLGR7 content.

3 Repertoire considerations

The current repertoire can now be simply derived from dotAsia by adding the following list:

- 18 characters from the Normalized Hanzi List for Common Use (NHCU)
- 2 HKSCS characters that were left out when processing HKIRC request
- 43 characters from JGP and KGP repertoire to complete variant set.

These last 43 characters are questionable because they seem to be only added to complete variant sets and are not used in Chinese, as far as the IP understands. If that is the case, these 43 characters should be only added as 'out-of-repertoire' (that is, having a reflexive mapping of "out-of-repertoire-var").

For the remainder, if we consider dotAsia to be an acceptable repertoire, the 18+2 extension makes sense. Some of the NHCU content might still be questioned as not being truly essential.

It should also be noted that the current Chinese repertoire seems larger than necessary, it is generally admitted that 2000-3000 Han ideographs are sufficient to read a Chinese newspaper, and knowing 8000 of these ideographs denotes a well-educated person, see http://www.bbc.co.uk/languages/chinese/real_chinese/mini_guides/characters/characters_howmany.shtml). Therefore, even for accommodating the need for a mix of simplified and traditional forms of these ideographs, the set could have been less than 10 000 in size.

Furthermore, having large repertoires for all CJK constituencies exacerbate coordination issues concerning variants sets. Larger repertoire increases the risk of common ideographs while the usage could be rare in one case and common in other cases, imposing unneeded variant derived restrictions on domain delegation for these constituencies.

Finally, the table of the 43 characters from JGP and KGP in pages 13 and 15 of the [Proposal] has many errors in its IICORE content. It should be corrected to be aligned with the same IICORE content exposed in the table shown in page 19-23 of the same document.

Integration Panel recommendations:

Consider reducing the size of the LGR and fix table content in page 13-14. In addition, the IP would like to remind the CJK community that the **modern** usage of each code point proposed in the LGR repertoire should be documented.

4 Variant considerations

4.1 General

As noted below, the level of required review varies vastly between the cases where variant sets have additional members compared to already deployed IDN tables, such as dotAsia, and the case where the variant sets are identical but the mappings are different. In the former case, the review is simply to ensure that the mapping makes sense given the nature of additions. In the latter case, one should

understand why CLGR7 would diverge from an existing IDN table which is already deployed and is addressing roughly the same repertoire.

Two tables in [Proposal] present various variant status, one in page 24-25, and another 26-27. By being listed piece by piece under code points and not singly under variant sets, they make the analysis unnecessarily difficult. For example, multiple entries refer to the same variant set and should be merged. In the first table (page 24-25), 5 of the 19 entries correspond to a single variant set. There are few errors (for example in the first entry, U+64E5 is clearly not a 'Simp', and the 'Trad' cell should not contain multiple values; it looks like the value U+39DB is missing and the other cell contents should be shifted one position right). In addition, that same table contains other errors (like mentioning U+3A5c belonging to CDNC). The other table (page 26-27) is more regular but again suffers by not presenting the data as variant sets.

Integration Panel recommendations:

In the [Proposal] any considerations about variant data should be presented in the form of variant sets, including all code points belonging to the same set in a single table in a format that could be similar to the one used in these considerations. This makes the analysis much easier to perform.

4.2 About J0 and K code points and out-of-repertoire variant issues

Normally, the inclusion of J-specific or K-specific code points in a Chinese LGR would appear to serve no purpose. From a repertoire perspective, it would only make sense if there was a requirement to apply for labels that combine these code points with some Chinese-only code points. Absent such a requirement, it is doubtful that the inclusion of these code points can be justified on repertoire considerations only. The Conservatism Principle demands that the repertoire selection be conservative – only the necessary code points should be included.

However, where these code points have variant relations with other code points that are in the CLGR repertoire, the issue becomes more interesting. Even if, under conservative design, a code point is only present in the Japanese LGR, for example, it might be possible to apply for a label that is seen, by Chinese users, as a variant of some other Chinese label. This cross-repertoire variant relation is similar to the cross-script variant issue in alphabetic scripts. In both cases, to allow for blocking the variant label, it is required to add the out-of-repertoire code point to the repertoire. If this is done, it normally is given a reflexive variant mapping (to itself) of type "out-of-repertoire-var" and variant mappings of type "blocked" to all code points that are variants of it in the repertoire.

Integration Panel recommendations:

In the case of CLGR7 it is now getting clear that the inclusion of J-specific or K-specific code points was done to handle such out-of-repertoire variant issues. Therefore, their mapping types should be updated so that they match the expected types for an out-of-repertoire code point as described above. If, instead, the GP desires them to be included as full members, the IP expects a documented justification for their inclusion as full members of the repertoire, based on their demonstrated use in Chinese establishing a requirement for support in IDNs.

4.3 Notation and explanation used in the description of the variant set differences

The following sections describe in considerable detail the differences between CLGR7 and dotAsia. Where the two LGR differ in some variant sets that otherwise overlap, the tables listing both variant sets are shown together and the header announces which one comes first or second. Actual differences are highlighted in red. Although the comparison was mostly mechanically generated there is some level of manual editing, therefore in case of discrepancies, the actual XML files are authoritative.

The format of each table follows the format used in the HTML-formatted LGR tables, such as used in LGR-1. The tables list each pair of variant mappings on one row. For each pair of code points, by convention, the lower code point is taken as the source of the mapping in the forward \rightarrow direction and information for the reverse direction \leftarrow is usually not listed separately. The variant mappings defined in an LGR are required to be symmetric, that is, both the forward and reverse mappings must be specified.

A mapping where source and target are the same is reflexive. Variant sets consisting of only a single reflexive mapping are not shown as a set. Instead, the variant type of the mapping is listed in the Variants column of the Repertoire by Code Point table. Reflexive mappings that are part of a larger set are indicated with a “ \equiv ”.

Where the type of both forward and reverse mappings are the same, a single value is given in the Type(s) column, otherwise the types for forward and reverse mapping are given in that order, as indicated by the arrows. The same applies to any comments.

In a properly specified LGR, all members of each variant set are variants of each other, a property called transitivity. Because of that, all variant sets are necessarily disjoint.

The variant sets are presented in increasing numerical order of source code points and target code points in the set, irrespective of which code point(s) have a difference in mapping between the LGRs being compared. This is to facilitate comparison with the original XML file (or HTML transcription).

When related sets from CLGR7 and dotAsia are presented together, the CLGR7 variant is always first.

In each table, the background color alternates whenever the source code point in the left column changes, thus grouping all mappings using the same source code point.

Some of the visual descriptions used in discussion of the CJK ideographs show the sources of these characters, as in this example for U+7ADA:

7ADA 立 117.5					
	GE-365F	H-8E56	T3-3323	JO-636C	K2-4F4D

Sources prefixed by G, H, T, J and K denote that a code point is sourced from China, Hong Kong, Taiwan, Japan and Korea, respectively. A second leading letter or digit designates a subset, for example “J0”, while the digits following the hyphen give a mapping to the specific source.

Note that G source GE (standing for GB16500-95) is a so-called ‘horizontal extension’. A horizontal extension provides an additional mapping for a code point, but does not establish usage on the same footing as ordinary sources. As such, the presence of a GE source value does not bring usage evidence for the source category. A code point with, for example, only a GE mapping and a JO mapping would normally be considered “Japanese-specific”.

4.4 About the status of the dotAsia variant set

The LGR document for CLGR7 claims in page 25 that the variants for the dotAsia IDN table were created as experimental for the HK characters (which constitute the clear majority of the additions to the repertoire from the original Dotcn/DotTw IDN tables) and that the *‘intent has always been to merge and make consistent with CGP table once it is integrated for root zone and gTLD purpose.’* Then in page 25 it is stated that *‘Moreover, for the 62 code points in dotAsia IDN table from Supplementary Plane, dotAsia agreed to set them all as INDEPENDENT characters in CGP Variants Mappings’.* The later statement implies that CJK Extension B Ideographs cannot have variant mappings, while currently in dotAsia, two of them (code points: U+282E2 and U+29D98) do have such a relationship.

These statements raise strong concerns because it implies a lack of stability for the variant set that could be very damaging for deployed domain names.

In addition, two tables (one in page 24-25 and the other in page 26-27) provide more details in the difference between CGR7 and dotAsia. The first table provides some justification for the difference, but only for 4 variants sets (out of the 111 that have differences). The second table states the CLGR7 version of the variant sets but does not provide any new rationale for these differences (beyond a consistency declaration mentioned in the first paragraph of this section 4.4).

A detailed feedback provided in section 5 of this document, using mainly Unihan as reference has shown that some of the differences can be explained by variant relationship shown in Unihan. The items in that section 5 where CLGR7 has good cases are the following: 5.2.4, 5.2.10, 5.2.42, 5.2.43, 5.2.45, 5.2.47, 5.3.3, and 5.3.4.

There are many cases where the recent additions to dotAsia do not participate in the traditional-simplified mapping (code points only available to the original label as denoted by a ‘r-neither’ variant mapping type) in CLGR7 while there were ‘r-both’ in dotAsia. No rationale is provided for that modification.

This pattern is repeated in the additions to dotAsia that participate in variant sets. Most of them are ‘r-neither’ as well.

In addition, there are many cases where Unihan does not provide any clues on how the variant mapping should be done, so the Integration Panel has no reference point to determine whether CLGR7 or dotAsia, or another reference source is preferable.

Integration Panel recommendation:

Addressing these issues could be done by providing references to all variant mappings using available sources such as UniHan, dotAsia, and any other relevant sources. Furthermore, these references should be stable (that is accessible for many years in the future). Finally, explain why most of the recent additions do not participate in traditional-simplified mapping.

4.5 Considerations on coordination between CJK GP concerning variants

The section 4.5 of the [Proposal] mentions in its first paragraph that *'some variants mappings are totally UNACCEPTABLE to one party'*; it should be noted that variant mappings are not shared among LGRs. While the integration process will create common variant sets among LGR sharing the same repertoire, the variant mappings are exclusive to each individual LGR before integration and the actions are performed on the individual LGRs. Therefore, the mention above is misleading.

Integration Panel recommendation:

The best strategy is to have both repertoires and variant sets as small as possible to decrease collision among LGRs sharing the same repertoire basis (such as CJK LGRs).

4.6 Consideration on multiple variant character mappings and multiple allocatable variants

Because of some earlier feedback from the Integration Panel on that topic, the section 4.6 of the [Proposal] goes into great length in exploring in how to reduce allocatable variants. One of the suggested option is to modify the reflexive mapping type and is being noted as 'suggested by IP' (middle of page 31). The IP did not suggest such a solution, but instead to reduce the number of multiple 'traditional' mappings.

Then the same section 4.6 explores a solution using multiple variant sub-type (ending by '-m'), but abandons it later in the document, while still leaving some remnants in section '5 Whole Label Evaluation Rules' which are not implemented in the repertoire part of the XML file [CLGR7]. If the proposed solution is not workable, it should not form part of the main document, but be relegated to an appendix as an explored but failed experiment.

5 Variant set differences by type of difference

5.1 Additional repertoire (CLGR7 adds code points not in dotAsia)

In these cases, the variant set in CLGR7 usually adds a single additional code point to a variant set defined in dotAsia, with the required mappings added; the other mappings retain the same variant types. These are expected extensions and are acceptable in most cases. They still need to be reviewed. In a few cases, there are additional differences; these are called out in the description of the variant sets in the following pages.

This summary table provides an overview of the impact of the 60 new code points on existing or new variant sets, showing the UniHan variant (when available), the IICORE value related to the new code

point (first letter is priority: A for high, C for low, other letters stand for Japanese (J), Chinese (G or T) or Korean (K or KP), and whether the new code point is part of the Normalized Hanzi list for Common Use (NHCUC). The entries that have no IICORE information are not part of that set.

No	New UCS	Unihan variant	IICORE	NHCUC
1	3960	8ADD	CK	
2*	3A5C	652C		
3	7ADA	4F47		Yes
4	4FAD	5118	AJ	
5	7E4B	7E6B	AJ	
6	51E6	8655	AJ	
7	524F	5231		Yes
8	5227	5226		Yes
9	56A2	56CA	AJ	
10*	58B5			
11	7E4A	7E96	ATJ	
12*	61F4	61FA	CJ	
13	6060	602A		Yes
14	6442	651D	AJ	
15	784F		AKP	
16	663B	6602	AKP	
17	9EB9		AJ	Yes
18	894D	96DC		Yes
19	685F	68E7	AJ	
20	685C	6AFB	AJ	
21	8262	6AA3		Yes
22	6E8C	6F51	AJ	
23	6D9C	7006	AJ	
24	731F	7375	AJ	
25	732F			Yes
26	74A2	7409		Yes
27	750E	78DA		Yes
28	754A	8015		Yes
29	9271	7926	AJ	
30	967A	96AA	AJ	
31	7A36		AKP	

No	New UCS	Unihan variant	IICORE	NHCUC
32	7B86		AJ	
33	7C14	7C11	CJ	
34	7D9A	7E8C	AJ	
35	81D3	81DF	AJ	
36	8133	8166	AJ	
37	984B	816E		Yes
38	8217	92EA	AJ	
39	839F		CJ	
40	83B5	83DF	CJ	
41	9D2C	9DAF	AJ	
42	86CD	87A2	AJ	
43	88B5	887D		Yes
44	8E99		CJ	
45	8F19	8F12		Yes
46	9039	9054	CJ	
47	91A4	91AC	AJ	
48	91C8	91CB	AJ	
49*	9421	9435	CJ	
50	945A	947D	CJ	Yes
51	96B2	9A2D	CJ	
52	9D8F	96DE	AJ	
53	9EBA	9EB5	AJ	
54	982C	9830	AJ	
55	98EE	98F2	AKP	
56	9A12	9A37	AJ	
57	9A13	9A57	AJ	
58	9A28	9A52	AJ	
59	9C2E	9CC1	CJ	
60	9D0E	9DD7	AJ	

* In some cases (2, 10, 12 and 49), the variant sets associated with these code points have other changes, these are mentioned in the following variant sets descriptions.

Some observations can be made from the summary table:

- 16 (14 of 15 NHCUC entries, plus U+3A5C and U+58B5) out of the 60 entries correspond to Chinese related additions. The other 44 cases correspond to Japanese or Korean related code points. It is not clear why these 44 new code points should be part of a Chinese LGR except as ‘out-of-repertoire-var’ type and mapping value ‘blocked’ for the variants in that code point entry (not presently the case).

- 8 entries have no Unihan variant defined in the Unihan database. In the absence of references for these entries, it is not possible to verify the validity of these additions into variant sets.

The following pages describe the 60 variant sets either modified or added to as result of these new 60 code points. The IP has conducted a preliminary evaluation to establish whether the chosen variant mappings appear reasonable, but would request the CGP to provide fuller documentation supporting the choices made.

1. This variant set has one added member U+3960.

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
3960	愔	3960	愔	≡ r-neither		identity
3960	愔	8ADD	諝	→ trad ← blocked		
3960	愔	8C1E	諝	→ simp ← blocked		
8ADD	諝	8ADD	諝	≡ r-trad		identity
8ADD	諝	8C1E	諝	→ simp ← trad		
8C1E	諝	8C1E	諝	≡ r-simp		identity

The code point U+3960 has G, T, J, and K source and is part of the IICORE set (value CK, meaning low priority, Korean usage).

3960
心 61.9
愔 愔 愔
G5-5436 T3-3B5F JA-2329
愔
K3-2554

Unihan kDefinition field indicates that this is a variant of U+8ADD 諝. As such the proposed mappings would be adequate if U+3960 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

2. This variant set has one added member U+3A5C (not in dotAsia.) from the HKSCS set. In addition, U+39DB and U+64E5 (both in dotAsia) are also included in the CLGR7 (1st) and are mapped differently from dotAsia (2nd). This case is a hybrid of this category (one code point added not in dotAsia) and the next category (two code points already in dotAsia but treated differently). The red highlighting in both tables reflects all differences between the two LGRs.

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
39DB	𢦏	39DB	𢦏	≡ r-simp		identity
39DB	𢦏	3A5C	𢦏	↔ blocked		
39DB	𢦏	63FD	攬	↔ blocked		
39DB	𢦏	64E5	𢦏	→ trad ← blocked		
39DB	𢦏	652C	攬	↔ blocked		
3A5C	𢦏	3A5C	𢦏	≡ r-both		identity
3A5C	𢦏	63FD	攬	↔ blocked		
3A5C	𢦏	64E5	𢦏	↔ blocked		
3A5C	𢦏	652C	攬	↔ blocked		
63FD	攬	63FD	攬	≡ r-simp		identity
63FD	攬	64E5	𢦏	↔ blocked		
63FD	攬	652C	攬	→ trad ← simp		
64E5	𢦏	64E5	𢦏	≡ r-both		identity

64E5	攬	652C	攬	↔	blocked		
652C	攬	652C	攬	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
39DB	攬	39DB	攬	≡	r-simp		identity
39DB	攬	63FD	攬	↔	blocked		
39DB	攬	64E5	攬	→	trad		
				←	simp		
39DB	攬	652C	攬	↔	blocked		
63FD	攬	63FD	攬	≡	r-simp		identity
63FD	攬	64E5	攬	↔	blocked		
63FD	攬	652C	攬	→	trad		
				←	simp		
64E5	攬	64E5	攬	≡	r-trad		identity
64E5	攬	652C	攬	↔	blocked		
652C	攬	652C	攬	≡	r-trad		identity

The code point U+3A5C has G, T, H, J, and V (Vietnam) source.

3A5C 手 64.14	攬 G5-4D25	攬 T3-5468	攬 JA-2348
	攬 V0-3875	攬 H-A078	

Unihan kDefinition field indicates that this is a variant of U+652C 攬. It also has semantic variants association with U+64E5 攬, and U+3A2B 攬(not in CLGR7) is listed as a simplified variant. While the mapping for U+3A5C is acceptable (and correspond to an earlier feedback from IP), there is no justification for changing the mapping for the pair (U+64E5, U+39DB). In fact, the table page 24-25 supports the dotAsia mapping for these 2 characters. Therefore, the following variant set definition is preferable:

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
39DB	𢦏	39DB	𢦏	≡	r-simp		identity
39DB	𢦏	3A5C	攔	↔	blocked		
39DB	𢦏	63FD	攬	↔	blocked		
39DB	𢦏	64E5	𢦏	→ ←	trad simp		
39DB	𢦏	652C	攬	↔	blocked		
3A5C	攔	3A5C	攔	≡	r-both		identity
3A5C	攔	63FD	攬	↔	blocked		
3A5C	攔	64E5	𢦏	↔	blocked		
3A5C	攔	652C	攬	↔	blocked		
63FD	攬	63FD	攬	≡	r-simp		identity
63FD	攬	64E5	𢦏	↔	blocked		
63FD	攬	652C	攬	→ ←	trad simp		
64E5	𢦏	64E5	𢦏	≡	r-trad		identity
64E5	𢦏	652C	攬	↔	blocked		
652C	攬	652C	攬	≡	r-trad		identity

3. This variant set has one added member U+7ADA.

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
4F2B	伫	4F2B	伫	≡ r-simp		identity
4F2B	伫	4F47	佇	→ trad ← simp		
4F2B	伫	7ADA	𠄎	→ blocked ← simp		
4F47	佇	4F47	佇	≡ r-trad		identity
4F47	佇	7ADA	𠄎	→ blocked ← trad		
7ADA	𠄎	7ADA	𠄎	≡ r-neither		identity

The code point U+7ADA has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

7ADA
立 117.5
𠄎 𠄎 𠄎 𠄎 𠄎
GE-365F H-8E56 T3-3323 J0-636C K2-4F4D

Unihan kSemanticVariant field indicates that this is a variant of U+4F47 佇. As such the proposed mappings are adequate.

4. This variant set has one added member U+4FAD. Note that it also needs a reflexive mapping “r-neither”.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
4FAD	𠄎	4FAD	𠄎	≡	r-neither		identity
4FAD	𠄎	5118	儘	→	trad		
				←	blocked		
4FAD	𠄎	5C3D	𠄎	→	simp		
				←	blocked		
4FAD	𠄎	76E1	盡	→	trad		
				←	blocked		
5118	儘	5118	儘	≡	r-trad		identity
5118	儘	5C3D	𠄎	→	simp		
				←	trad		
5118	儘	76E1	盡	↔	blocked		
5C3D	𠄎	5C3D	𠄎	≡	r-simp		identity
5C3D	𠄎	76E1	盡	→	trad		
				←	simp		
76E1	盡	76E1	盡	≡	r-trad		identity

The code point U+4FAD has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

4FAD
人 9.6

𠄎
GE-2168

𠄎
JO-4B79

Unihan kZVariant field indicates that this is a variant of U+5118 儘. As such the proposed mappings would be adequate if U+4FAD was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

5. This variant set has one added member U+7E4B.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
4FC2	係	4FC2	係	≡	r-trad		identity
4FC2	係	7CFB	系	→	simp		
				←	trad		
4FC2	係	7E4B	繫	↔	blocked		
4FC2	係	7E6B	繫	↔	blocked		
7CFB	系	7CFB	系	≡	r-both		identity
7CFB	系	7E4B	繫	→	blocked		
				←	simp		
7CFB	系	7E6B	繫	→	trad		
				←	simp		
7E4B	繫	7E6B	繫	→	trad		
				←	blocked		
7E4B	繫	7E4B	繫	≡	r-neither		identity
7E6B	繫	7E6B	繫	≡	r-trad		identity

The code point U+7E4B has G, J, and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

7E4B 係 120.13 繫 GE-3859


繫 繫 J0-3752 K2-5331


Unihan kZvariant field indicates that this is a variant of U+7E6B 繫. As such the proposed mappings are adequate if U+7E4B was required for Chinese usage. Note that U+4FC2, U+7CFB, and U+7E4B have J0 sources. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

6. This variant set has one added member U+51E6.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
51E6	処	51E6	処	≡	r-neither		identity
51E6	処	5904	处	→	simp		
				←	blocked		
51E6	処	8655	處	→	trad		
				←	blocked		
5904	处	5904	处	≡	r-simp		identity
5904	处	8655	處	→	trad		
				←	simp		
8655	處	8655	處	≡	r-trad		identity

The code point U+51E6 has G, J, and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

51E6
几 16.3

GE-227E

 
J0-3D68 K2-244F

UniHan kZVariant field indicates that this is a variant of U+8655 處. It is also seen as a Semantic variant of U+458F (not in this table). As such the proposed mappings may still be adequate if U+51E6 was required for Chinese usage. Note that U+51E6 and U+8655 have J0 sources. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

7. This variant set has one added member U+524F.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
521B	创	521B	创	≡	r-simp		identity
521B	创	5231	𠄎	→	blocked		
				←	simp		
521B	创	524F	𠄎	→	blocked		
				←	simp		
521B	创	5259	𠄎	→	blocked		
				←	simp		
521B	创	5275	創	→	trad		
				←	simp		
5231	𠄎	5231	𠄎	≡	r-trad		identity
5231	𠄎	524F	𠄎	↔	blocked		
5231	𠄎	5259	𠄎	↔	blocked		
5231	𠄎	5275	創	↔	blocked		
524F	𠄎	524F	𠄎	≡	r-neither		identity
524F	𠄎	5259	𠄎	↔	blocked		
524F	𠄎	5275	創	→	trad		
				←	blocked		
5259	𠄎	5259	𠄎	≡	r-neither		identity

5259	𠄎	5275	創	→ trad ← blocked		
5275	創	5275	創	≡ r-trad		identity

The code point U+524F has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

524F
刀 18.7

𠄎 𠄎 𠄎 𠄎 𠄎

GE-2339 H-87BC T3-2B43 J0-516C K2-246B

Unihan kZVariant field indicates that this is a variant of U+5231 𠄎. As such, the mapping U+524F→U+5231 should be ‘trad’ and U+524F F→U+5275 should be ‘blocked’. See table below for a modified variant set:

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
521B	創	521B	創	≡ r-simp		identity
521B	創	5231	𠄎	→ blocked ← simp		
521B	創	524F	𠄎	→ blocked ← simp		
521B	創	5259	𠄎	→ blocked ← simp		
521B	創	5275	創	→ trad ← simp		
5231	𠄎	5231	𠄎	≡ r-trad		identity
5231	𠄎	524F	𠄎	→ blocked ← trad		

5231	𠄎	5259	𠄎	↔	blocked		
5231	𠄎	5275	創	↔	blocked		
524F	𠄎	524F	𠄎	≡	r-neither		identity
524F	𠄎	5259	𠄎	↔	blocked		
524F	𠄎	5275	創	↔	blocked		
5259	𠄎	5259	𠄎	≡	r-neither		identity
5259	𠄎	5275	創	→	trad		
				←	blocked		
5275	創	5275	創	≡	r-trad		identity

This feedback was provided in the context of CLGR6 review and was not acted on (except for the addition of the identity mapping for U+524F and U+5259) on and not commented either.

8. This variant set has one added member U+5227.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5226	刂	5226	刂	≡	r-neither		identity
5226	刂	5227	刂	↔	blocked		
5226	刂	523C	刂	↔	blocked		
5226	刂	52AB	刂	→	both		
				←	blocked		
5227	刂	5227	刂	≡	r-neither		identity
5227	刂	523C	刂	↔	blocked		
5227	刂	52AB	刂	→	both		
				←	blocked		
523C	刂	523C	刂	≡	r-neither		identity
523C	刂	52AB	刂	→	both		
				←	blocked		
52AB	刂	52AB	刂	≡	r-both		identity

The code point U+5227 has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

5227
刀 18.5
 刂 刂 刂 刂 刂
GE-2330 H-89A6 T3-245D J0-5167 K1-582E

Unihan kSemanticVariant indicates field that this is a variant of U+5226 刂 and U+52AB 刂. As such the proposed mappings are adequate. Note that U+5227 and U+52AB have J0 sources.

9. This new variant set is created by combining the new U+56A2 with the existing U+56CA in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant.)

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
56A2	囊	56A2	囊	≡	r-neither		identity
56A2	囊	56CA	囊	→	both		
				←	blocked		
56CA	囊	56CA	囊	≡	r-both		identity

The code point U+56A2 has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).



Unihan kZVariant field indicates that this is a variant of U+56CA 囊. As such the proposed mappings may still be adequate if U+56A2 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

10. This variant set has one added member U+58B5 (not in dotAsia.) from the HKSCS set. In addition, U+58DC (in dotAsia) is also included in the CLGR7 (1st) and is mapped differently from dotAsia (2nd). This case is a hybrid of this category (one code point added not in dotAsia) and the next category (one code point already in dotAsia but treated differently). The dotAsia table does not include U+57EE and U+58B0 in this variant set (both are singleton reflexive variants of type ‘r-both’), while the CLGR7 table add them as ‘blocked’ variants of all other members. The red highlighting in both table reflects all differences between the two LGRs.
The CLGR7 is an 8 members variant set.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
575B	坛	575B	坛	≡	r-simp		identity
575B	坛	57EE	焮	↔	blocked		
575B	坛	58B0	壇	↔	blocked		
575B	坛	58B5	壇	↔	blocked		
575B	坛	58C7	壇	→ ←	trad simp		
575B	坛	58DC	壇	↔	blocked		
575B	坛	7F48	罈	→ ←	trad simp		
575B	坛	7F4E	罈	→ ←	blocked simp		
57EE	焮	57EE	焮	≡	r-both		identity
57EE	焮	58B0	壇	↔	blocked		
57EE	焮	58B5	壇	↔	blocked		

57EE	焱	58C7	壇	↔	blocked		
57EE	焱	58DC	壘	↔	blocked		
57EE	焱	7F48	罈	↔	blocked		
57EE	焱	7F4E	罍	↔	blocked		
58B0	壇	58B0	壇	≡	r-both		identity
58B0	壇	58B5	壘	↔	blocked		
58B0	壇	58C7	壇	↔	blocked		
58B0	壇	58DC	壘	↔	blocked		
58B0	壇	7F48	罈	↔	blocked		
58B0	壇	7F4E	罍	↔	blocked		
58B5	壘	58B5	壘	≡	r-both		identity
58B5	壘	58C7	壇	↔	blocked		
58B5	壘	58DC	壘	↔	blocked		
58B5	壘	7F48	罈	↔	blocked		
58B5	壘	7F4E	罍	↔	blocked		
58C7	壇	58C7	壇	≡	r-trad		identity
58C7	壇	58DC	壘	↔	blocked		

58C7	壇	7F48	罫	↔	blocked		
58C7	壇	7F4E	𪗗	→ ←	blocked trad		
58DC	壘	58DC	壘	≡	r-both		identity
58DC	壘	7F48	罫	↔	blocked		
58DC	壘	7F4E	𪗗	↔	blocked		
7F48	罫	7F48	罫	≡	r-trad		identity
7F48	罫	7F4E	𪗗	↔	blocked		
7F4E	𪗗	7F4E	𪗗	≡	r-neither		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
575B	坛	575B	坛	≡	r-simp		identity
575B	坛	58C7	壇	→ ←	trad simp		
575B	坛	58DC	壘	→ ←	blocked simp		
575B	坛	7F48	罫	→ ←	trad simp		
575B	坛	7F4E	𪗗	→ ←	blocked simp		
58C7	壇	58C7	壇	≡	r-trad		identity

58C7	壇	58DC	壘	→ blocked ← trad		
58C7	壇	7F48	罈	↔ blocked		
58C7	壇	7F4E	罍	→ blocked ← trad		
58DC	壘	58DC	壘	≡ r-both		identity
58DC	壘	7F48	罈	↔ blocked		
58DC	壘	7F4E	罍	↔ blocked		
7F48	罈	7F48	罈	≡ r-trad		identity
7F48	罈	7F4E	罍	↔ blocked		
7F4E	罍	7F4E	罍	≡ r-neither		identity

The new CLGR7 corresponds to the recommendation that was made by IP on earlier feedback and as such the result is satisfactory. However, the fact that this creates another 8 member variant set with multiple allocatable mappings is not a situation that is reassuring.

11. This variant set has one added member U+7E4A.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5B45	織	5B45	織	≡	r-trad		identity
5B45	織	7E34	緯	↔	blocked		
5B45	織	7E4A	織	↔	blocked		
5B45	織	7E8E	織	↔	blocked		
5B45	織	7E96	織	↔	blocked		
5B45	織	7EA4	纤	→ ←	simp blocked		
7E34	緯	7E34	緯	≡	r-trad		identity
7E34	緯	7E4A	織	↔	blocked		
7E34	緯	7E8E	織	↔	blocked		
7E34	緯	7E96	織	↔	blocked		
7E34	緯	7EA4	纤	→ ←	simp blocked		
7E4A	織	7E4A	織	≡	r-neither		identity
7E4A	織	7E8E	織	↔	blocked		
7E4A	織	7E96	織	→ ←	trad blocked		
7E4A	織	7EA4	纤	→	simp		

				← blocked		
7E8E	織	7E8E	織	≡ r-neither		identity
7E8E	織	7E96	織	→ trad ← blocked		
7E8E	織	7EA4	纤	→ simp ← blocked		
7E96	織	7E96	織	≡ r-trad		identity
7E96	織	7EA4	纤	→ simp ← trad		
7EA4	纤	7EA4	纤	≡ r-simp		identity

The code point U+7E4A has G, J and K sources and is part of the IICORE set (value ATJ, meaning high priority, Chinese-Traditional and Japanese usage).

7E4A 系 120.11 織 GE-3858

織 織 JO-4121 K2-5330

Unihan kZVariant field indicates that this is a variant of U+7E96 織. As such the proposed mappings may still be adequate if U+7EA4 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

12. This variant set has one added member U+61F4. Note that it also needs a reflexive mapping “r-neither”. In addition, the mappings between U+61FA and U+5FCF have different type assignments between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5FCF	忤	5FCF	忤	≡	r-both		identity
5FCF	忤	61F4	懺	→	blocked		
				←	simp		
5FCF	忤	61FA	懺	→	trad		
				←	simp		
61F4	懺	61F4	懺	≡	r-neither		identity
61F4	懺	61FA	懺	→	trad		
				←	blocked		
61FA	懺	61FA	懺	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5FCF	忤	5FCF	忤	≡	r-both		identity
5FCF	忤	61FA	懺	→	blocked		
				←	simp		
61FA	懺	61FA	懺	≡	r-trad		identity

The code point U+61F4 has G, J and K sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

61F4
心 61.15

GE-2A6C

 
J0-5873 K2-337A

Unihan kZVariant field indicates that this is a variant of U+61FA 懺. As such the proposed mappings may still be adequate if U+61F4 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’. The mappings between U+61FA and U+5FCF in CLGR7 seem to follow the Unihan model (in it U+61FA is the Traditional Variant of 5FCF).

13. This new variant set is created by combining the added U+6060 with the existing U+602A in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
602A	怪	602A	怪	≡	r-both		identity
602A	怪	6060	恠	→	blocked		
				←	simp		
6060	恠	6060	恠	≡	r-trad		identity

The code point U+6060 has G, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

6060
心 61.6

恠 恠 恠 恠

GE-2A21 T3-2C5F J0-575F K1-5A26

Unihan kSemanticVariant field indicates that this is a variant of U+602A 怪. As such the proposed mappings are adequate.

14. This variant set has one added member U+6442.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6315	扨	6315	扨	≡	r-trad		identity
6315	扨	6442	扨	↔	blocked		
6315	扨	6444	撮	→	simp		
				←	blocked		
6315	扨	651D	撮	↔	blocked		
6442	扨	6442	扨	≡	r-neither		identity
6442	扨	6444	撮	→	simp		
				←	blocked		
6442	扨	651D	撮	→	trad		
				←	blocked		
6444	撮	6444	撮	≡	r-simp		identity
6444	撮	651D	撮	→	trad		
				←	simp		
651D	撮	651D	撮	≡	r-trad		identity

The code point U+6442 has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

6442
手 64.10 扨 撮
GE-2B6A JO-405D

Unihan kZVariant field indicates that this is a variant of U+651D 撮. As such the proposed mappings may still be adequate if U+6442 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

15. This variant set has one added member U+784F.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
63C5	𦉑	63C5	𦉑	≡	r-trad		identity
63C5	𦉑	7814	研	→	simp		
				←	blocked		
63C5	𦉑	784F	研	↔	blocked		
7814	研	7814	研	≡	r-both		identity
7814	研	784F	研	→	blocked		
				←	both		
784F	研	784F	研	≡	r-neither		identity

The code point U+784F has G, T, J and K sources and is part of the IICORE set (value AKP, meaning high priority, Korean usage).

784F
石 112.6

研 GE-3534

研 T3-3761

研 J3-7923


研 K0-665A

Unihan does not define any variant for this character, therefore no conclusion can be made based on its IICORE classification. If instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

16. This new variant set is created by combining the new U+663B with the existing U+6602 in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant.)

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6602	昂	6602	昂	≡	r-both		identity
6602	昂	663B	昂	→	blocked		
				←	both		
663B	昂	663B	昂	≡	r-neither		identity

The code point U+663B has G, T, and K sources and is part of the IICORE set (value AKP, meaning high priority, Korean usage).

663B
日 72.5


Unihan kSemanticVariant field indicates that this is a variant of U+6602 昂. As such the proposed mappings may still be adequate if U+663B was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

17. This variant set has one added member U+9EB9.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
66F2	曲	66F2	曲	≡	r-both		identity
66F2	曲	9EAF	麴	→ ←	blocked both		
66F2	曲	9EB4	麴	→ ←	trad simp		
66F2	曲	9EB9	麴	→ ←	blocked both		
9EAF	麴	9EAF	麴	≡	r-neither		identity
9EAF	麴	9EB4	麴	↔	blocked		
9EAF	麴	9EB9	麴	↔	blocked		
9EB4	麴	9EB4	麴	≡	r-trad		identity
9EB4	麴	9EB9	麴	→ ←	blocked trad		
9EB9	麴	9EB9	麴	≡	r-neither		identity

The code point U+9EB9 has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9EB9 麥 199:8 麴 GE-486C

9EB9 麴 JO-396D

Unihan does not define any variant for this character, therefore no conclusion can be made. Based on its IICORE classification it is not clear why U+9EB9 is included in the NHCUC set and therefore needed in CLGR7 context.

18. This variant set has one added member U+894D.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6742	杂	6742	杂	≡	r-simp		identity
6742	杂	894D	襍	→	blocked		
				←	simp		
6742	杂	96D1	雜	→	blocked		
				←	simp		
6742	杂	96DC	雜	→	trad		
				←	simp		
6742	杂	96E5	糴	→	blocked		
				←	simp		
894D	襍	894D	襍	≡	r-neither		identity
894D	襍	96D1	雜	↔	blocked		
894D	襍	96DC	雜	→	trad		
				←	blocked		
894D	襍	96E5	糴	↔	blocked		
96D1	雜	96D1	雜	≡	r-neither		identity
96D1	雜	96DC	雜	→	trad		
				←	blocked		
96D1	雜	96E5	糴	↔	blocked		

96DC	雜	96DC	雜	≡	r-trad		identity
96DC	雜	96E5	𪗇	↔	blocked		
96E5	𪗇	96E5	𪗇	≡	r-trad		identity

The code point U+894D has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

894D
衣 145.12
 襪 襪 襪 襪 襪
GE-3D3E H-95A7 T3-587C J0-7037 K1-6B3D

Unihan kSemanticVariant field indicates that this is a variant of U+96DC 雜. As such the proposed mappings are adequate.

19. This variant set has one added member U+685F.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6808	棧	6808	棧	≡	r-simp		identity
6808	棧	685F	棧	→	blocked		
				←	simp		
6808	棧	68E7	棧	→	trad		
				←	simp		
6808	棧	8F4F	𪗇	→	blocked		
				←	simp		
685F	棧	685F	棧	≡	r-neither		identity
685F	棧	68E7	棧	→	trad		
				←	blocked		
685F	棧	8F4F	𪗇	↔	blocked		
68E7	棧	68E7	棧	≡	r-trad		identity
68E7	棧	8F4F	𪗇	↔	blocked		
8F4F	𪗇	8F4F	𪗇	≡	r-trad		identity

The code point U+685F has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

685F
木 75.6
棧
GE-2D7B

棧
JO-3B37

UniHan kZVariant field indicates that this is a variant of U+68E7 棧. As such the proposed mappings may still be adequate if U+685F was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

20. This variant set has one added member U+685C.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
685C	桜	685C	桜	≡	r-neither		identity
685C	桜	6A31	櫻	→	simp		
				←	blocked		
685C	桜	6AFB	櫻	→	trad		
				←	blocked		
6A31	櫻	6A31	櫻	≡	r-simp		identity
6A31	櫻	6AFB	櫻	→	trad		
				←	simp		
6AFB	櫻	6AFB	櫻	≡	r-trad		identity

The code point U+685C has G, J and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

685C 桜
木 75.6
GE-2D78

桜 櫻
JO-3A79 K2-3A6C

Unihan kZVariant field indicates that this is a variant of U+6AFB 櫻. As such the proposed mappings may still be adequate if U+685C was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

21. This variant set has one added member U+8262. Note that it also needs a reflexive mapping “r-neither”.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6A2F	檣	6A2F	檣	≡	r-simp		identity
6A2F	檣	6AA3	檣	→	trad		
				←	simp		
6A2F	檣	8262	艦	→	blocked		
				←	simp		
6AA3	檣	6AA3	檣	≡	r-trad		identity
6AA3	檣	8262	艦	→	blocked		
				←	trad		
8262	艦	8262	艦	≡	r-neither		identity

The code point U+8262 has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

8262 舟 137.13 艦 艦 艦 艦 艦
GE-3A2F H-956B T3-5A7A J0-6760 K2-5669

Unihan kSemanticVariant field indicates that this is a variant of U+6AA3 檣. As such the proposed mappings are adequate.

22. This variant set has one added member U+6E8C.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6CFC	泼	6CFC	泼	≡	r-simp		identity
6CFC	泼	6E8C	澆	→	blocked		
				←	simp		
6CFC	泼	6F51	潑	→	trad		
				←	simp		
6E8C	澆	6E8C	澆	≡	r-neither		identity
6E8C	澆	6F51	潑	→	trad		
				←	blocked		
6F51	潑	6F51	潑	≡	r-trad		identity

The code point U+685C has G, J and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

6E8C
水 85.9
澆
GE-3065

澆 澆
JO-482E K2-4158

Unihan kZVariant field indicates that this is a variant of U+6F51 潑. As such the proposed mappings may still be adequate if U+6E8C was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

23. This variant set has one added member U+6D9C.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6D9C	澆	6D9C	澆	≡	r-neither		identity
6D9C	澆	6E0E	澆	→	simp		
				←	blocked		
6D9C	澆	7006	澆	→	trad		
				←	blocked		
6E0E	澆	6E0E	澆	≡	r-simp		identity
6E0E	澆	7006	澆	→	trad		
				←	simp		
7006	澆	7006	澆	≡	r-trad		identity

The code point U+6D9C has G, J and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

6D9C
水 85.7
澆
GE-303A

澆 澆
J0-4642 K2-4066

Unihan kZVariant field indicates that this is a variant of U+7006 澆. As such the proposed mappings may still be adequate if U+6D9C was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

24. This variant set has one added member U+731F.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
730E	猎	730E	猎	≡	r-simp		identity, reflexive
730E	猎	731F	獵	→	blocked		
				←	simp		
730E	猎	7375	獵	→	trad		
				←	simp		
731F	獵	731F	獵	≡	r-neither		identity
731F	獵	7375	獵	→	trad		
				←	blocked		
7375	獵	7375	獵	≡	r-trad		identity, reflexive

The code point U+731F has G, J and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

731F
犬 94.8
獵
GE-3323

獵 獵
J0-4E44 K2-4676

UniHan kZVariant field indicates that this is a variant of U+7375 獵. As such the proposed mappings may still be adequate if U+731F was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

25. This new variant set is created by combining the new U+732F with the existing U+8C92 in a single variant set. (In dotAsia the latter is a singleton reflexive 'r-both' variant).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
732F	𤝵	732F	𤝵	≡	r-both		identity, reflexive
732F	𤝵	8C92	𤝵	↔	blocked		
8C92	𤝵	8C92	𤝵	≡	r-both		identity, reflexive

The code point U+8262 has G, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

732F
犬 94.9
𤝵
G5-467E

𤝵 𤝵 𤝵
T3-3D2D J0-604E K2-467E

Unihan does not define any variant for this character, therefore no conclusion can be made. Given the sources, it is not yet clear why U+732F is part of the NHCU set.

26. This variant set has one added member U+74A2. In addition, the mapping between U+7409 and U+7460 is different between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7409	琉	7409	琉	≡	r-both		identity
7409	琉	7460	瑠	→	blocked		
				←	simp		
7409	琉	74A2	璠	→	blocked		
				←	simp		
7460	瑠	7460	瑠	≡	r-trad		identity
7460	瑠	74A2	璠	↔	blocked		
74A2	璠	74A2	璠	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7409	琉	7409	琉	≡	r-both		identity
7409	琉	7460	瑠	↔	blocked		
7460	瑠	7460	瑠	≡	r-both		identity

The code point U+74A2 has G, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

74A2
 玉 96.13
 璠 GE-3353 璠 璠 璠
 T3-516B J0-6065 K2-484E

Unihan kSemanticVariant field indicates that this is a variant of U+7409 琉. Unihan does not bring any clarification to the discrepancy. Furthermore, given the sources, it is not yet clear why U+732F is part of the NHCUs set.

27. This variant set has one added member U+750E.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
750E	𪗇	750E	𪗇	≡	r-neither		identity
750E	𪗇	7816	砖	→	simp		
				←	blocked		
750E	𪗇	78DA	磚	→	trad		
				←	blocked		
7816	砖	7816	砖	≡	r-simp		identity
7816	砖	78DA	磚	→	trad		
				←	simp		
78DA	磚	78DA	磚	≡	r-trad		identity

The code point U+750E has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

750E
瓦 98.11

𪗇 𪗇 𪗇 𪗇 𪗇

GE-336D H-FEA9 T3-516E J0-612F K1-6C24

Unihan kSemanticVariant field indicates that this is a variant of U+78DA 磚. As such the proposed mappings are adequate.

28. This new variant set is created by combining the new U+754A with the existing U+8015 in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
754A	畊	754A	畊	≡	r-trad		identity
754A	畊	8015	耕	→	simp		
				←	blocked		
8015	耕	8015	耕	≡	r-both		identity

The code point U+754A has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

754A
田 102.4
 畊 畊 畊 畊 畊
GE-3379 H-FEB2 T3-2E28 J0-6139 K1-5854

Unihan kSemanticVariant field indicates that this is a variant of U+8015 耕. As such the proposed mappings are adequate.

29. This variant set has one added member U+9271.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
77FF	𠩺	77FF	𠩺	≡	r-simp		identity
77FF	𠩺	7926	礦	→	trad		
				←	simp		
77FF	𠩺	9271	鉷	→	blocked		
				←	simp		
77FF	𠩺	945B	鑛	→	blocked		
				←	simp		
7926	礦	7926	礦	≡	r-trad		identity
7926	礦	9271	鉷	→	blocked		
				←	trad		
7926	礦	945B	鑛	→	blocked		
				←	trad		
9271	鉷	9271	鉷	≡	r-neither		
9271	鉷	945B	鑛	↔	blocked		
954B	銳	954B	銳	≡	r-neither		identity

The code point U+9271 has G, J and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9271
金 167.5
鉷
GE-4167

鉷 鉷
J0-395B K2-6849

Unihan kZVariant field indicates that this is a variant of U+7926 礦. As such the proposed mappings may still be adequate if U+9271 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

30. This variant set has one added member U+967A.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7877	硷	7877	硷	≡	r-simp		identity
7877	硷	78B1	碱	↔	blocked		
7877	硷	7906	硷	→ ←	blocked simp		
7877	硷	9669	险	↔	blocked		
7877	硷	967A	險	↔	blocked		
7877	硷	96AA	險	↔	blocked		
7877	硷	9E7C	鹵	→ ←	trad simp		
78B1	碱	78B1	碱	≡	r-simp		identity
78B1	碱	7906	硷	↔	blocked		
78B1	碱	9669	险	↔	blocked		
78B1	碱	967A	險	↔	blocked		
78B1	碱	96AA	險	↔	blocked		
78B1	碱	9E7C	鹵	→ ←	trad blocked		
7906	硷	7906	硷	≡	r-neither		identity
7906	硷	9669	险	↔	blocked		

7906	險	967A	險	↔	blocked		
7906	險	96AA	險	→	trad		
				←	blocked		
7906	險	9E7C	險	↔	blocked		
9669	險	9669	險	≡	r-simp		identity
9669	險	967A	險	→	blocked		
				←	simp		
9669	險	96AA	險	→	trad		
				←	simp		
9669	險	9E7C	險	↔	blocked		
967A	險	967A	險	≡	r-neither		identity
967A	險	96AA	險	→	trad		
				←	blocked		
967A	險	9E7C	險	↔	blocked		
96AA	險	96AA	險	≡	r-trad		identity
96AA	險	9E7C	險	↔	blocked		
9E7C	險	9E7C	險	≡	r-trad		identity

The code point U+967A has G, J and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

967A
草 170.8
險
GE-442D

險 險
J0-3831 K2-8B6F

Unihan kZVariant field indicates that this is a variant of U+96AA 險. As such the proposed mappings may still be adequate if U+967A was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

31. This new variant set is created by combining the new U+7A36 with the existing U+7A22 in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant.)

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7A22	穧	7A22	穧	≡	r-both		identity
7A22	穧	7A36	穧	→	blocked		
				←	both		
7A36	穧	7A36	穧	≡	r-neither		identity

The code point U+7A36 has G, T, J and K sources and is part of the IICORE set (value AKP, meaning high priority, Korean usage).

7A36
禾 115.10

穧
GE-3634

穧
T4-4B6B

穧
K0-6973

Unihan does not define any variant for this character, therefore no conclusion can be made. As such the proposed mappings may still be adequate if U+7A36 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

32. This new variant set is created by combining the new U+7B86 with the existing U+7BE6 in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant.)

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7B86	筧	7B86	筧	≡	r-neither		identity
7B86	筧	7BE6	筧	→	both		
				←	blocked		
7BE6	筧	7BE6	筧	≡	r-both		identity

The code point U+7B86 has G, T, J and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

7B86
竹 118.8

筧 GE-3724

筧 T3-4860

筧 JO-4A4F

筧 K2-5030

Unihan does not define any variant for this character, therefore no conclusion can be made. As such the proposed mappings may still be adequate if U+7B86 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

33. This variant set has one added member U+7C14.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7C11	𦵱	7C11	𦵱	≡	r-trad		identity
7C11	𦵱	7C14	𦵱	→	blocked		
				←	trad		
7C11	𦵱	84D1	𦵱	→	simp		
				←	trad		
7C14	𦵱	7C14	𦵱	≡	r-neither		identity
7C14	𦵱	84D1	𦵱	→	simp		
				←	blocked		
84D1	𦵱	84D1	𦵱	≡	r-simp		identity

The code point U+7C14 has G, T and J sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

7C14 竹 118.11 𦵱
 GE-373F

𦵱 𦵱
 TF-5B5F J0-6443

Unihan kZVariant field indicates that this is a variant of U+7C11 𦵱. As such the proposed mappings may still be adequate if U+7C14 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

34. This variant set has one added member U+7D9A.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7D9A	統	7D9A	統	≡	r-neither		identity
7D9A	統	7E8C	續	→	trad		
				←	blocked		
7D9A	統	7EED	续	→	simp		
				←	blocked		
7E8C	續	7E8C	續	≡	r-trad		identity
7E8C	續	7EED	续	→	simp		
				←	trad		
7EED	续	7EED	续	≡	r-simp		identity

The code point U+7D9A has G, J and K sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

7D9A
糸 120.7 統
GE-383A

統 統
J0-4233 K2-5250

Unihan kZVariant field indicates that this is a variant of U+7E8C 續. As such the proposed mappings may still be adequate if U+7D9A was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

35. This variant set has one added member U+81D3.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
810F	𩇑	810F	𩇑	≡	r-simp		identity
810F	𩇑	81D3	𩇑	→	blocked		
				←	simp		
810F	𩇑	81DF	𩇑	→	trad		
				←	simp		
810F	𩇑	9AD2	𩇑	→	trad		
				←	simp		
81D3	𩇑	81D3	𩇑	≡	r-neither		identity
81D3	𩇑	81DF	𩇑	→	trad		
				←	blocked		
81D3	𩇑	9AD2	𩇑	↔	blocked		
81DF	𩇑	81DF	𩇑	≡	r-trad		identity
81DF	𩇑	9AD2	𩇑	↔	blocked		
9AD2	𩇑	9AD2	𩇑	≡	r-trad		identity

The code point U+81D3 has G, J and KP sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

81D3
肉 130.14 𩇑
GE-3974

𩇑
J0-4221

UniHan kZVariant field indicates that this is a variant of U+81DF 𩇑. As such the proposed mappings may still be adequate if U+81D3 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

36. This variant set has one added member U+8133.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
8111	腦	8111	腦	≡	r-simp		identity
8111	腦	8133	腦	→	blocked		
				←	simp		
8111	腦	8166	腦	→	trad		
				←	simp		
8133	腦	8133	腦	≡	r-neither		identity
8133	腦	8166	腦	→	trad		
				←	blocked		
8166	腦	8166	腦	≡	r-trad		identity

The code point U+8133 has G, J and KP sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

8133
肉 130.7

GE-3963


JO-473E

Unihan kZVariant field indicates that this is a variant of U+8166 腦. As such the proposed mappings may still be adequate if U+8133 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

37. This new variant set is created by combining the new U+984B with the existing U+816E in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
816E	腮	816E	腮	≡	r-both		identity
816E	腮	984B	顛	→	blocked		
				←	simp		
984B	顛	984B	顛	≡	r-trad		identity

The code point U+984B has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

984B
頁 181.9

顛 顛 顛 顛 顛

GE-453C H-90F8 T3-5959 J0-707B K1-6569

Unihan kSemanticVariant field indicates that this is a variant of U+816E 腮. As such the proposed mappings are adequate. (Although the point of making U+816E a simplified mapping for U+984B could be investigated.)

38. This variant set has one added member U+8217.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
8216	舖	8216	舖	≡	r-trad		identity
8216	舖	8217	舖	→	blocked		
				←	trad		
8216	舖	92EA	舖	↔	blocked		
8216	舖	94FA	铺	→	simp		
				←	trad		
8217	舖	8217	舖	≡	r-neither		identity
8217	舖	92EA	舖	↔	blocked		
8217	舖	94FA	铺	→	simp		
				←	blocked		
92EA	舖	92EA	舖	≡	r-trad		identity
92EA	舖	94FA	铺	→	simp		
				←	trad		
94FA	铺	94FA	铺	≡	r-simp		identity

The code point U+8217 has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

8217
舌 135.9
GE-3A24

舖

舖
JO-4A5E

Unihan kZVariant field indicates that this is a variant of U+92EA 舖. As such the proposed mappings may still be adequate if U+8217 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'. Note that U+8216, U+8217, and U+92EA have J0 sources.

39. This new variant set is created by combining the new U+839F with the existing U+83E1 in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant.)

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
839F	蒼	839F	蒼	≡	r-neither		identity
839F	蒼	83E1	菡	→	both		
				←	blocked		
83E1	菡	83E1	菡	≡	r-both		identity

The code point U+839F has G, T, J and K sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

839F 蒼 蒼 蒼 蒼
艸 140.7 G3-692F T3-3857 J0-6831 K1-727E

Unihan does not define any variant for this character, therefore no conclusion can be made. As such the proposed mappings may still be adequate if U+839F was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

40. This new variant set is created by combining the new U+83B5 with the existing U+83DF in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant.)

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
83B5	菟	83B5	菟	≡	r-neither		identity
83B5	菟	83DF	菟	→	both		
				←	blocked		
83DF	菟	83DF	菟	≡	r-both		identity

The code point U+83B5 has G and J sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

83B5
艸 140.7
 菟
GE-3A5A

菟
J0-6839

Unihan kZVariant field indicates that this is a variant of U+83DF 菟. As such the proposed mappings may still be adequate if U+83B5 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

41. This variant set has one added member U+9D2C.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
83BA	鸛	83BA	鸛	≡	r-simp		identity
83BA	鸛	9D2C	鶯	→	blocked		
				←	simp		
83BA	鸛	9DAF	鶯	→	trad		
				←	simp		
9D2C	鶯	9D2C	鶯	≡	r-neither		identity
9D2C	鶯	9DAF	鶯	→	trad		
				←	blocked		
9DAF	鶯	9DAF	鶯	≡	r-trad		identity

The code point U+9D2C has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9D2C
鳥 196.5
鶯
GE-4763

鶯
JO-3229

Unihan kZVariant field indicates that this is a variant of U+9DAF 鶯. As such the proposed mappings may still be adequate if U+9D2C was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

42. This variant set has one added member U+86CD.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
8424	螢	8424	螢	≡	r-simp		identity
8424	螢	86CD	螢	→	blocked		
				←	simp		
8424	螢	87A2	螢	→	trad		
				←	simp		
86CD	螢	86CD	螢	≡	r-neither		identity
86CD	螢	87A2	螢	→	trad		
				←	blocked		
87A2	螢	87A2	螢	≡	r-trad		identity

The code point U+86CD has G, J and KP sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

86CD
虫 142.5

GE-3C2A


J0-3756

Unihan kZVariant field indicates that this is a variant of U+87A2 螢. As such the proposed mappings may still be adequate if U+9D2C was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

43. This new variant set is created by combining the new U+88B5 with the existing U+887D in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
887D	衿	887D	衿	≡	r-both		identity
887D	衿	88B5	衿	→	blocked		
				←	simp		
88B5	衿	88B5	衿	≡	r-trad		identity

The code point U+88B5 has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

88B5
衿 145.6
衿 衿 衿 衿 衿
GE-3D29 H-8FC4 T3-3E72 J0-6A54 K2-5E4C

Unihan kSemanticVariant field indicates that this is a variant of U+887D 衿. As such the proposed mappings are adequate.

44. This variant set has one added member U+8E99.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
8E8F	躡	8E8F	躡	≡	r-simp		identity
8E8F	躡	8E99	躡	→	blocked		
				←	simp		
8E8F	躡	8EAA	躡	→	trad		
				←	simp		
8E99	躡	8E99	躡	≡	r-neither		identity
8E99	躡	8EAA	躡	→	trad		
				←	blocked		
8EAA	躡	8EAA	躡	≡	r-trad		identity

The code point U+8E99 has G, T, J and K sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

8E99
足 157.16

躡 躡 躡 躡

G5-6F39 T3-604B J0-6D38 K1-5F6D

Unihan does not define any variant for this character, therefore no conclusion can be made. As such the proposed mappings may still be adequate if U+8E99 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

45. This variant set has one added member U+8F19.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
8F12	輒	8F12	輒	≡	r-trad		identity
8F12	輒	8F19	輒	→	blocked		
				←	trad		
8F12	輒	8F84	輒	→	simp		
				←	trad		
8F19	輒	8F19	輒	≡	r-neither		identity
8F19	輒	8F84	輒	→	simp		
				←	blocked		
8F84	輒	8F84	輒	≡	r-simp		identity

The code point U+8F19 has G, H, T, J, and K sources and is part of the Normalized Hanzi list for Common Use.

8F19
車 159.8
輒 輒 輒 輒 輒
GE-3F73 H-904C T3-4F36 J0-6D4D K2-6527

UniHan kSemanticVariant field indicates that this is a variant of U+8F12 輒. As such the proposed mappings are adequate.

46. This variant set has one added member U+9039.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
8FBE	达	8FBE	达	≡	r-simp		identity
8FBE	达	8FD6	达	→	blocked		
				←	simp		
8FBE	达	9039	達	→	blocked		
				←	simp		
8FBE	达	9054	達	→	trad		
				←	simp		
8FD6	达	8FD6	达	≡	r-trad		identity
8FD6	达	9039	達	↔	blocked		
8FD6	达	9054	達	↔	blocked		
9039	達	9039	達	≡	r-neither		identity
9039	達	9054	達	→	trad		
				←	blocked		
9054	達	9054	達	≡	r-trad		identity

The code point U+9039 has G, J and KP sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

9039 達 達
飛 162.8 GE-4043 JO-6D7D

Unihan kZVariant field indicates that this is a variant of U+9054 達. As such the proposed mappings may still be adequate if U+9039 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

47. This variant set has one added member U+91A4.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9171	醬	9171	醬	≡	r-simp		identity
9171	醬	91A4	醬	→	blocked		
				←	simp		
9171	醬	91AC	醬	→	trad		
				←	simp		
91A4	醬	91A4	醬	≡	r-neither		identity
91A4	醬	91AC	醬	→	trad		
				←	blocked		
91AC	醬	91AC	醬	≡	r-trad		identity

The code point U+91A4 has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

91A4	將	將
西 164.10	西	西
	GE-4124	JO-3E5F

UniHan kZVariant field indicates that this is a variant of U+91AC 醬. As such the proposed mappings may still be adequate if U+91A4 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

48. This variant set has one added member U+91C8.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
91C8	𪗇	91C8	𪗇	≡	r-neither		identity
91C8	𪗇	91CA	𪗈	→	simp		
				←	blocked		
91C8	𪗇	91CB	𪗉	→	trad		
				←	blocked		
91CA	𪗈	91CA	𪗈	≡	r-simp		identity
91CA	𪗈	91CB	𪗉	→	trad		
				←	simp		
91CB	𪗉	91CB	𪗉	≡	r-trad		identity

The code point U+91C8 has G, J, K and KP sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

91C8
采 165.4
𪗇
GE-412C

𪗇 𪗇
JO-3C61 K2-6757

Unihan kZVariant field indicates that this is a variant of U+91CB 釋. As such the proposed mappings may still be adequate if U+91C8 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

49. This variant set has one added member U+9421 (not in dotAsia.) Note that it also needs a reflexive mapping “r-neither”. In addition, U+9244 (in dotAsia) was also included in the CLGR7 (1st) and is mapped differently from dotAsia (2nd). This case is a hybrid of this category (one code point added not in dotAsia) and the next category (one point already in dotAsia but treated differently).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
9244	鉄	9244	鉄	≡ r-neither		identity
9244	鉄	9295	𨮑	↔ blocked		
9244	鉄	9421	鐵	↔ blocked		
9244	鉄	9435	鐵	→ trad ← blocked		
9244	鉄	94C1	铁	→ simp ← blocked		
9295	𨮑	9295	𨮑	≡ r-trad		identity
9295	𨮑	9421	鐵	↔ blocked		
9295	𨮑	9435	鐵	↔ blocked		
9295	𨮑	94C1	铁	→ simp ← blocked		
9421	鐵	9421	鐵	≡ r-neither		identity
9421	鐵	9435	鐵	→ trad ← blocked		
9421	鐵	94C1	铁	→ simp ← blocked		
9435	鐵	9435	鐵	≡ r-trad		identity

9435	鐵	94C1	铁	→ simp ← trad		
94C1	铁	94C1	铁	≡ r-simp		identity
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
9244	鐵	9244	鐵	≡ r-trad		identity
9244	鐵	9295	鍊	↔ blocked		
9244	鐵	9435	鐵	↔ blocked		
9244	鐵	94C1	铁	→ simp ← trad		
9295	鍊	9295	鍊	≡ r-trad		identity
9295	鍊	9435	鐵	↔ blocked		
9295	鍊	94C1	铁	→ simp ← blocked		
9435	鐵	9435	鐵	≡ r-trad		identity
9435	鐵	94C1	铁	→ simp ← trad		
94C1	铁	94C1	铁	≡ r-simp		identity

The code point U+9421 has G, J and K sources and is part of the IICORE set (value CJ, meaning low

priority, Japanese usage).

9421
金 167.12
鐵

鐵 鐵
JO-6F45 K2-6A5A

Unihan kZVariant field indicates that this is a variant of U+9435 鐵. As such the proposed mappings may still be adequate if U+9421 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'. The differences concerning U+9244 mappings are not addressed at this point (Unihan Semantic Variant mappings between U+9244, U+9295, and U+9435).

50. This variant set has one added member U+945A.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9246	鈔	9246	鈔	≡	r-trad		identity
9246	鈔	9409	鐸	↔	blocked		
9246	鈔	945A	鑽	↔	blocked		
9246	鈔	947D	鑽	↔	blocked		
9246	鈔	94BB	钻	→	simp		
				←	blocked		
9409	鐸	9409	鐸	≡	r-trad		identity
9409	鐸	945A	鑽	↔	blocked		
9409	鐸	947D	鑽	↔	blocked		
9409	鐸	94BB	钻	→	simp		
				←	blocked		
945A	鑽	945A	鑽	≡	r-neither		identity
945A	鑽	947D	鑽	→	trad		
				←	blocked		
945A	鑽	94BB	钻	→	simp		
				←	blocked		
947D	鑽	947D	鑽	≡	r-trad		identity
947D	鑽	94BB	钻	→	simp		
				←	trad		

94BB	钻	94BB	钻	≡	r-simp		identity
------	---	------	---	---	--------	--	----------

The code point U+945A has G, T, J, K, and KP sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

945A
金 167.15 鑽 GE-4349

鑽 T3-6055

鑽 J0-6F54

鑽 K2-6A75

Unihan kSemanticVariant field indicates that this is a variant of U+947D 鑽. As such the proposed mappings may still be adequate if U+945A was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

51. This variant set has one added member U+96B2.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
96B2	隳	96B2	隳	≡	r-neither		identity
96B2	隳	9A2D	鷲	→	trad		
				←	blocked		
96B2	隳	9A98	鷲	→	simp		
				←	blocked		
9A2D	鷲	9A2D	鷲	≡	r-trad		identity
9A2D	鷲	9A98	鷲	→	simp		
				←	trad		
9A98	鷲	9A98	鷲	≡	r-simp		identity

The code point U+96B2 has G, T, and J sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

96B2
阜 170.14

GE-443D T3-5671 J0-702D

Unihan kSemanticVariant field indicates that this is a variant of U+9A2D 鷲. As such the proposed mappings may still be adequate if U+96B2 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

52. This variant set has one added member U+9D8F.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
96DE	雞	96DE	雞	≡	r-trad		identity
96DE	雞	9CEE	𪗇	→ ←	blocked trad		
96DE	雞	9D8F	鷄	→ ←	blocked trad		
96DE	雞	9DC4	鷄	→ ←	blocked trad		
96DE	雞	9E21	鸡	→ ←	simp trad		
9CEE	𪗇	9CEE	𪗇	≡	r-neither		identity
9CEE	𪗇	9D8F	鷄	↔	blocked		
9CEE	𪗇	9DC4	鷄	↔	blocked		
9CEE	𪗇	9E21	鸡	→ ←	simp blocked		
9D8F	鷄	9D8F	鷄	≡	r-neither		identity
9D8F	鷄	9DC4	鷄	↔	blocked		
9D8F	鷄	9E21	鸡	→ ←	simp blocked		
9DC4	鷄	9DC4	鷄	≡	r-neither		identity

9DC4	鷄	9E21	鸡	→ simp		
				← blocked		
9E21	鸡	9E21	鸡	≡ r-simp		identity

The code point U+9D8F has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9D8F 鳥 196.8	鷄	鷄
	GE-4825	JO-375C

UniHan kSemanticVariant field indicates that this is a variant of U+96DE 雞. As such the proposed mappings may still be adequate if U+9D8F was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

53. This variant set has one added member U+9EBA.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9762	面	9762	面	≡	r-both		identity
9762	面	9EAA	麩	→	blocked		
				←	simp		
9762	面	9EB5	麵	→	trad		
				←	simp		
9762	面	9EBA	麵	→	blocked		
				←	simp		
9EAA	麩	9EAA	麩	≡	r-neither		identity
9EAA	麩	9EB5	麵	→	trad		
				←	blocked		
9EAA	麩	9EBA	麵	↔	blocked		
9EB5	麵	9EB5	麵	≡	r-trad		identity
9EB5	麵	9EBA	麵	→	blocked		
				←	trad		
9EBA	麵	9EBA	麵	≡	r-neither		identity

The code point U+9EBA has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9EBA	麵	麵
麦 199:9		
GE-486D		JO-4C4D

Unihan kZVariant field indicates that this is a variant of U+9EB5 麵. As such the proposed mappings may still be adequate if U+9EBA was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

54. This variant set has one added member U+982C.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
982C	頰	982C	頰	≡	r-neither		identity
982C	頰	9830	頰	→	trad		
				←	blocked		
982C	頰	988A	頰	→	simp		
				←	blocked		
9830	頰	9830	頰	≡	r-trad		identity
9830	頰	988A	頰	→	simp		
				←	trad		
988A	頰	988A	頰	≡	r-simp		identity

The code point U+982C has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

982C
頁 181.6
頰
GE-4532

頰
J0-4B4B

Unihan kZvariant field indicates that this is a variant of U+9830 頰. As such the proposed mappings may still be adequate if U+982C was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

55. This variant set has one added member U+98EE.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
98EE	飲	98EE	飲	≡	r-neither		identity
98EE	飲	98F2	飲	→	trad		
				←	blocked		
98EE	飲	996E	饮	→	simp		
				←	blocked		
98F2	飲	98F2	飲	≡	r-trad		identity
98F2	飲	996E	饮	→	simp		
				←	trad		
996E	饮	996E	饮	≡	r-simp		identity

The code point U+98EE has G and J sources and is part of the IICORE set (value AKP meaning high priority, Korean usage).

98EE
食 184.4
飲
GE-454E

飲 飲
JO-5D3B KO-6B66

Unihan kZvariant field indicates that this is a variant of U+98F2 飲. As such the proposed mappings may still be adequate if U+98EE was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

56. This variant set has one added member U+9A12.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9A12	騷	9A12	騷	≡	r-neither		identity
9A12	騷	9A37	騷	→	trad		
				←	blocked		
9A12	騷	9A9A	騷	→	simp		
				←	blocked		
9A37	騷	9A37	騷	≡	r-trad		identity
9A37	騷	9A9A	騷	→	simp		
				←	trad		
9A9A	騷	9A9A	騷	≡	r-simp		identity

The code point U+9A12 has G, J and KP sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9A12 騷 騷
馬 187.8 GE-462C J0-417B

Unihan kZVariant field indicates that this is a variant of U+9A37 騷. As such the proposed mappings may still be adequate if U+9A12 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

57. This variant set has one added member U+9A13. Note that it also needs a reflexive mapping “r-neither”.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9A13	驗	9A13	驗	≡	r-neither		identity
9A13	驗	9A57	驗	→	trad		
				←	blocked		
9A13	驗	9A8C	驗	→	simp		
				←	blocked		
9A57	驗	9A57	驗	≡	r-trad		identity
9A57	驗	9A8C	驗	→	simp		
				←	trad		
9A8C	驗	9A8C	驗	≡	r-simp		identity

The code point U+9A13 has G, J, K, and KP sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9A13
馬 187.8 驗 驗 驗
GE-462D J0-3833 K2-6F2C

Unihan kZVariant field indicates that this is a variant of U+9A57 驗. As such the proposed mappings may still be adequate if U+9A13 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

58. This new variant set is created by combining the new U+9A28 with the existing U+9A52 in a single variant set. (In dotAsia the latter is a singleton reflexive ‘r-both’ variant).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9A28	𩇑	9A28	𩇑	≡	r-neither		identity
9A28	𩇑	9A52	𩇒	→	both		
			𩇒	←	blocked		
9A52	𩇒	9A52	𩇒	≡	r-both		identity, reflexive

The code point U+9A28 has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9A28 馬 187.9	𩇑 GE-4635	𩇑 JO-424D
-----------------	--------------	--------------

Unihan kZVariant field indicates that this is a variant of U+9A52 𩇒. As such the proposed mappings may still be adequate if U+9A28 was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

59. This variant set has one added member U+9C2E.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9C1B	鯧	9C1B	鯧	≡	r-trad		identity
9C1B	鯧	9C2E	鯧	↔	blocked		
9C1B	鯧	9CC1	鯧	→	simp		
				←	trad		
9C2E	鯧	9C2E	鯧	≡	r-trad		identity
9C2E	鯧	9CC1	鯧	→	simp		
				←	blocked		
9CC1	鯧	9CC1	鯧	≡	r-simp		identity

The code point U+9C2E has G, T, J, K and KP sources and is part of the IICORE set (value CJ, meaning low priority, Japanese usage).

9C2E
魚 195.10

鯧 GE-4742

鯧 T3-5F2A

鯧 J0-725B

鯧 K1-6846

Unihan kZVariant field indicates that this is a variant of U+9CC1 鯧. As such the proposed mappings may still be adequate if U+9C2E was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all ‘blocked’.

60. This variant set has one added member U+9D0E.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9D0E	𪗇	9D0E	𪗇	≡	r-neither		identity
9D0E	𪗇	9DD7	𪗇	→	trad		
				←	blocked		
9D0E	𪗇	9E25	𪗇	→	simp		
				←	blocked		
9DD7	𪗇	9DD7	𪗇	≡	r-trad		identity
9DD7	𪗇	9E25	𪗇	→	simp		
				←	trad		
9E25	𪗇	9E25	𪗇	≡	r-simp		identity

The code point U+9D0E has G and J sources and is part of the IICORE set (value AJ, meaning high priority, Japanese usage).

9D0E 𪗇 𪗇
鳥 196.4 GE-475B JO-322A

Unihan kZVariant field indicates that this is a variant of U+9DD7 𪗇. As such the proposed mappings may still be adequate if U+9D0E was required for Chinese usage. However, if instead its justification is purely from an integration scenario, its mapping to other code points should be all 'blocked'.

5.2 Code points added to CLGR7 through dotAsia but treated differently

The variant sets in this section have code points added from the dotAsia repertoire, but the variant mappings chosen differ from those used in the original dotAsia set. To show the differences, variant sets may be listed twice, once for CLGR7 and once for dotAsia.

The IP has not yet evaluated the merits of these different choices and would request the GCP to provide documentation of the rationale, background, references etc. that motivate the chosen mappings.

1. The code point U+3A18 was included in CLGR7 because of its membership in IICORE but it has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd)

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
3A18	搯	3A18	搯	≡	r-trad		identity
3A18	搯	64E4	擻	→	simp		
				←	blocked		
64E4	擻	64E4	擻	≡	r-both		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
3A18	搯	3A18	搯	≡	r-both		identity
3A18	搯	64E4	擻	↔	blocked		
64E4	擻	64E4	擻	≡	r-both		identity

The simplified mapping between U+3A18 and U+64E4 in CLG7 is not supported by UniHan and looks doubtful.

2. The code point U+3A52 was included in CLGR7 because of its membership in IICORE but is treated differently between CLGR7 (table follows) and dotAsia (where it is a singleton reflexive variant of ‘r-both’)

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
3A52	𢮒	3A52	𢮒	≡ r-trad		identity
3A52	𢮒	64D2	𢮓	→ simp ← blocked		
64D2	𢮓	64D2	𢮓	≡ r-both		identity

The simplified mapping between U+3A52 and U+64D2 in CLGR7 is not supported by Unihan and looks doubtful. Unihan kDefinition field for U+3A52 indicates that this is a variant of U+64D2 but without simplified mapping. In Unihan U+64D2 has itself a semantic variant relationship with U+6366 𢮓, not supported by either CLGR7 or dotAsia.

3. In dotAsia, U+4882 and U+282E2 have a variant relationship. In CLGR7, U+4882 is a singleton reflexive variant of type ‘r-both’.

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
4882	𣪗	4882	𣪗	≡ r-simp		identity
4882	𣪗	282E2	𣪘	→ trad ← simp		
282E2	𣪘	282E2	𣪘	≡ r-trad		identity

Unihan has a simplified mapping between U+282E2 and U+4882. Not having it in CLGR7 is in error.

4. The code point U+4C7D was included in CLGR7 because of its membership in IICORE and the code point U+4C9D was added because of its GS reference (Singapore source). In the CLGR7 (1st), U+4C7D and U+4C9D are part of a larger variant set. In dotAsia (2nd), U+4C7D and U+4C9D form a smaller variant set together with compatible values; similarly, the pair U+9BE7 and U+9CB3 forms another variant set with compatible values.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
4C7D	𩺑	4C7D	𩺑	≡	r-trad		identity
4C7D	𩺑	4C9D	𩺒	→	simp		
				←	trad		
4C7D	𩺑	9BE7	𩺓	↔	blocked		
4C7D	𩺑	9CB3	𩺓	↔	blocked		
4C9D	𩺒	4C9D	𩺒	≡	r-simp		identity
4C9D	𩺒	9BE7	𩺓	↔	blocked		
4C9D	𩺒	9CB3	𩺓	↔	blocked		
9BE7	𩺓	9BE7	𩺓	≡	r-trad		identity
9BE7	𩺓	9CB3	𩺓	→	simp		
				←	trad		
9CB3	𩺓	9CB3	𩺓	≡	r-simp		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
4C7D	𩺑	4C7D	𩺑	≡	r-trad		identity

4C7D	鯨	4C9D	鯨	→ simp		
				← trad		
4C9D	鯨	4C9D	鯨	≡ r-simp		identity

Other dotAsia variant set:

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
9BE7	鯨	9BE7	鯨	≡	r-trad		identity
9BE7	鯨	9CB3	鯨	→ simp			
				← trad			
9CB3	鯨	9CB3	鯨	≡	r-simp		identity

The Unihan entries for U+47CD, U+4C9D, U+9BE7 and U+9CB3 supports the CLGR7 mappings.

5. In dotAsia, U+4C9E and U+29D98 have a variant relationship. In CLGR7, U+4C9E is a singleton reflexive variant of type 'r-both'.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
4C9E	鯨	4C9E	鯨	≡	r-simp		identity
4C9E	鯨	29D98	鯨	→	trad		
				←	simp		
29D98	鯨	29D98	鯨	≡	r-trad		identity

Unihan has a simplified mapping between U+29D98 and U+4C9E. Not having it in CLGR7 is in error.

6. The code point U+53DA was included in CLGR7 because of its membership in IICORE but is treated differently between CLGR7 (1st) and dotAsia (2nd). In one case, it is a variant of U+6BB5, in the other a variant of U+5047. The code points U+6BB5 and U+5047 are members of both CLGR7 and dotAsia with 'r-both' mapping.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
53DA	段	53DA	段	≡	r-both		identity
53DA	段	6BB5	段	↔	blocked		
6BB5	段	6BB5	段	≡	r-both		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5047	假	5047	假	≡	r-both		identity
5047	假	53DA	段	↔	blocked		
53DA	段	53DA	段	≡	r-both		identity

Unihan does not bring any clarification either way. This needs further investigation.

7. The code point U+701E was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between dotAsia (table follows) and CLGR7 (in that case it is just 'r-both').

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
51C0	淨	51C0	淨	≡ r-simp		identity
51C0	淨	51C8	淨	→ blocked ← simp		
51C0	淨	6D44	淨	→ blocked ← simp		
51C0	淨	6DE8	淨	→ trad ← simp		
51C0	淨	701E	澗	→ blocked ← simp		
51C8	淨	51C8	淨	≡ r-trad		identity
51C8	淨	6D44	淨	↔ blocked		
51C8	淨	6DE8	淨	↔ blocked		
51C8	淨	701E	澗	↔ blocked		
6D44	淨	6D44	淨	≡ r-neither		identity
6D44	淨	6DE8	淨	→ trad ← blocked		
6D44	淨	701E	澗	↔ blocked		
6DE8	淨	6DE8	淨	≡ r-trad		identity

6DE8	淨	701E	淨	→ blocked		
				← trad		
701E	淨	701E	淨	≡ r-neither		identity

Unihan does not bring any clarification either way. This needs further investigation.

8. The code point U+681E was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
520A	刊	520A	刊	≡	r-both		identity
520A	刊	520B	刊	→	blocked		
				←	both		
520A	刊	681E	栞	→	blocked		
				←	simp		
520B	刊	520B	刊	≡	r-neither		identity
520B	刊	681E	栞	↔	blocked		
681E	栞	681E	栞	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
520A	刊	520A	刊	≡	r-both		identity
520A	刊	520B	刊	→	blocked		
				←	both		
520A	刊	681E	栞	→	blocked		
				←	both		
520B	刊	520B	刊	≡	r-neither		identity
520B	刊	681E	栞	↔	blocked		
681E	栞	681E	栞	≡	r-neither		identity

Mappings between U+520A and U+520B are supported by UniHan, but U+681E is not mentioned. This needs further investigation.

9. The code point U+52A4 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between dotAsia (table follows) and CLGR7 (where it is a singleton reflexive variant of type 'r-both')

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
52A4	𠂔	52A4	𠂔	≡ r-both		identity
52A4	𠂔	52B2	𠂔	↔ blocked		
52A4	𠂔	52C1	𠂔	↔ blocked		
52B2	𠂔	52B2	𠂔	≡ r-simp		identity
52B2	𠂔	52C1	𠂔	→ trad ← simp		
52C1	𠂔	52C1	𠂔	≡ r-trad		identity

Unihan does not bring any clarification either way. This needs further investigation.

10. The code point U+767A was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (table follows) and dotAsia (not shown). The only difference is the type of reflexive mapping for U+767A, 'r-neither' for CLGR7 and 'r-trad' for dotAsia.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
53D1	发	53D1	发	≡	r-simp		identity
53D1	发	5F42	𠄎	→	blocked		
				←	simp		
53D1	发	767A	𠄎	→	blocked		
				←	simp		
53D1	发	767C	發	→	trad		
				←	simp		
53D1	发	9AEA	髮	→	blocked		
				←	simp		
53D1	发	9AEE	髮	→	trad		
				←	simp		
5F42	𠄎	5F42	𠄎	≡	r-neither		identity
5F42	𠄎	767A	𠄎	↔	blocked		
5F42	𠄎	767C	發	→	trad		
				←	blocked		
5F42	𠄎	9AEA	髮	↔	blocked		
5F42	𠄎	9AEE	髮	↔	blocked		

767A	𦍋	767A	𦍋	≡	r-neither		identity
767A	𦍋	767C	𦍋	→	trad		
				←	blocked		
767A	𦍋	9AEA	𦍋	↔	blocked		
767A	𦍋	9AEE	𦍋	↔	blocked		
767C	𦍋	767C	𦍋	≡	r-trad		identity
767C	𦍋	9AEA	𦍋	↔	blocked		
767C	𦍋	9AEE	𦍋	↔	blocked		
9AEA	𦍋	9AEA	𦍋	≡	r-neither		identity
9AEA	𦍋	9AEE	𦍋	→	trad		
				←	blocked		
9AEE	𦍋	9AEE	𦍋	≡	r-trad		identity

Unihan would tend to support the CLGR7 mapping, because only U+767C and U+9AEA are traditional mapping for U+53D1. Furthermore, U+767A is shown as a kZvariant of U+767C in Unihan. But this should be confirmed.

11. The code point U+8117 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
543B	吻	543B	吻	≡	r-both		identity
543B	吻	5445	咬	→	blocked		
				←	simp		
543B	吻	5461	眠	→	blocked		
				←	simp		
543B	吻	8117	膾	→	blocked		
				←	both		
5445	咬	5445	咬	≡	r-trad		identity
5445	咬	5461	眠	↔	blocked		
5445	咬	8117	膾	↔	blocked		
5461	眠	5461	眠	≡	r-trad		identity
5461	眠	8117	膾	↔	blocked		
8117	膾	8117	膾	≡	r-neither		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
543B	吻	543B	吻	≡	r-both		identity
543B	吻	5445	咬	→	blocked		
				←	simp		
543B	吻	5461	眠	→	blocked		

				← simp		
543B	吻	8117	脞	↔ blocked		
5445	咬	5445	咬	≡ r-trad		identity
5445	咬	5461	眠	↔ blocked		
5445	咬	8117	脞	↔ blocked		
5461	眠	5461	眠	≡ r-trad		identity
5461	眠	8117	脞	↔ blocked		
8117	脞	8117	脞	≡ r-both		identity

Unihan does not bring any clarification; this needs further study.

12. The code point U+8597 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
56ED	园	56ED	园	≡ r-simp		identity
56ED	园	5712	園	→ trad		
				← simp		
56ED	园	8597	菌	→ blocked		
				← simp		
5712	園	5712	園	≡ r-trad		identity
5712	園	8597	菌	→ blocked		
				← trad		
8597	菌	8597	菌	≡ r-neither		identity
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
56ED	园	56ED	园	≡ r-simp		identity
56ED	园	5712	園	→ trad		
				← simp		
56ED	园	8597	菌	→ blocked		
				← simp		
5712	園	5712	園	≡ r-trad		identity
5712	園	8597	菌	↔ blocked		
8597	菌	8597	菌	≡ r-trad		identity

Unihan does not bring any clarification; this needs further study.

13. The code point U+73E4 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (table follows) and dotAsia (where it is a singleton reflexive variant of type 'r-both').

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5B9D	宝	5B9D	宝	≡	r-simp		identity
5B9D	宝	5BF3	寶	→	blocked		
				←	simp		
5B9D	宝	5BF6	寶	→	trad		
				←	simp		
5B9D	宝	73E4	珎	→	blocked		
				←	simp		
5BF3	寶	5BF3	寶	≡	r-neither		identity
5BF3	寶	5BF6	寶	→	trad		
				←	blocked		
5BF3	寶	73E4	珎	↔	blocked		
5BF6	寶	5BF6	寶	≡	r-trad		identity
5BF6	寶	73E4	珎	→	blocked		
				←	trad		
73E4	珎	73E4	珎	≡	r-neither		identity

Unihan does not bring any clarification; this needs further study.

14. The code point U+7B92 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5E1A	𦍋	5E1A	𦍋	≡	r-both		identity
5E1A	𦍋	7B92	𦍌	→	blocked		
				←	both		
5E1A	𦍋	83F7	𦍋	→	blocked		
				←	both		
7B92	𦍌	7B92	𦍌	≡	r-neither		identity
7B92	𦍌	83F7	𦍋	↔	blocked		
83F7	𦍋	83F7	𦍋	≡	r-neither		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5E1A	𦍋	5E1A	𦍋	≡	r-both		identity
5E1A	𦍋	7B92	𦍌	↔	blocked		
5E1A	𦍋	83F7	𦍋	→	blocked		
				←	both		
7B92	𦍌	7B92	𦍌	≡	r-both		identity
7B92	𦍌	83F7	𦍋	↔	blocked		
83F7	𦍋	83F7	𦍋	≡	r-neither		identity

Unihan indicates that U+5E1A is a semantic variant of U+7B92. Therefore, the correlation is established, but not the type of mapping and why there is a difference between CLGR7 and dotAsia.

15. The code point U+6335 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5F04	弄	5F04	弄	≡	r-both		identity
5F04	弄	6335	拏	→	blocked		
				←	simp		
5F04	弄	8856	𢀆	→	blocked		
				←	simp		
6335	拏	6335	拏	≡	r-trad		identity
6335	拏	8856	𢀆	↔	blocked		
8856	𢀆	8856	𢀆	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5F04	弄	5F04	弄	≡	r-both		identity
5F04	弄	6335	拏	↔	blocked		
5F04	弄	8856	𢀆	→	blocked		
				←	simp		
6335	拏	6335	拏	≡	r-both		identity
6335	拏	8856	𢀆	↔	blocked		
8856	𢀆	8856	𢀆	≡	r-trad		identity

Unihan does not bring any clarification; this needs further study.

16. The code point U+60E3 was included in CLGR7 (1st) because of its membership in IICORE but is treated differently. In dotAsia, it is part of another variant set with U+63D4 (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
603B	总	603B	总	≡	r-simp		identity
603B	总	60E3	惣	→	blocked		
				←	simp		
603B	总	6374	惣	→	blocked		
				←	simp		
603B	总	6403	总	→	blocked		
				←	simp		
603B	总	6460	惣	→	blocked		
				←	simp		
603B	总	7DCF	総	→	blocked		
				←	simp		
603B	总	7E02	總	→	blocked		
				←	simp		
603B	总	7E3D	總	→	trad		
				←	simp		
60E3	惣	60E3	惣	≡	r-neither		identity
60E3	惣	6374	惣	↔	blocked		
60E3	惣	6403	总	↔	blocked		
60E3	惣	6460	惣	↔	blocked		

60E3	惣	7DCF	総	↔	blocked		
60E3	惣	7E02	総	↔	blocked		
60E3	惣	7E3D	総	→	trad		
				←	blocked		
6374	惣	6374	惣	≡	r-neither		identity
6374	惣	6403	总	↔	blocked		
6374	惣	6460	摠	↔	blocked		
6374	惣	7DCF	総	↔	blocked		
6374	惣	7E02	総	↔	blocked		
6374	惣	7E3D	総	→	trad		
				←	blocked		
6403	总	6403	总	≡	r-neither		identity
6403	总	6460	摠	↔	blocked		
6403	总	7DCF	総	↔	blocked		
6403	总	7E02	総	↔	blocked		
6403	总	7E3D	總	→	trad		
				←	blocked		
6460	摠	6460	摠	≡	r-trad		identity
6460	摠	7DCF	総	↔	blocked		

6460	摠	7E02	總	↔	blocked		
6460	摠	7E3D	總	↔	blocked		
7DCF	總	7DCF	總	≡	r-neither		identity
7DCF	總	7E02	總	↔	blocked		
7DCF	總	7E3D	總	→	trad		
				←	blocked		
7E02	總	7E02	總	≡	r-neither		identity
7E02	總	7E3D	總	→	trad		
				←	blocked		
7E3D	總	7E3D	總	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
60E3	惣	60E3	惣	≡	r-trad		identity
60E3	惣	63D4	惣	→	simp		
				←	blocked		
63D4	惣	63D4	惣	≡	r-both		identity

Unihan does not bring any clarification; this needs further study.

17. The code point U+617D was included in CLGR7 because of its membership in IICORE but CLGR7 (1st) and dotAsia (2nd). In one case, it is associated with U+617C, in the other with U+621A.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
617C	感	617C	感	≡	r-both		identity
617C	感	617D	憾	↔	blocked		
617D	憾	617D	憾	≡	r-both		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
617D	憾	617D	憾	≡	r-both		identity
617D	憾	621A	戚	↔	blocked		
621A	戚	621A	戚	≡	r-both		identity

In Unihan U+617C is a kSemanticVariant of U+617D and vice versa. There is no such association with U+621A. In consequence, the CLGR7 mapping is probably preferable.

18. The code point U+656D was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and addotAsia (2nd). Note that U+656D is mostly used in a Korean context. UniHan does not bring clarification.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
626C	扬	626C	扬	≡	r-simp		identity
626C	扬	63DA	揚	→	trad		
				←	simp		
626C	扬	656D	𪗇	→	blocked		
				←	simp		
626C	扬	98BA	颺	↔	blocked		
626C	扬	98CF	颺	↔	blocked		
63DA	揚	63DA	揚	≡	r-trad		identity
63DA	揚	656D	𪗇	↔	blocked		
63DA	揚	98BA	颺	↔	blocked		
63DA	揚	98CF	颺	↔	blocked		
656D	𪗇	656D	𪗇	≡	r-trad		identity
656D	𪗇	98BA	颺	↔	blocked		
656D	𪗇	98CF	颺	↔	blocked		
98BA	颺	98BA	颺	≡	r-trad		identity
98BA	颺	98CF	颺	→	simp		
				←	trad		

98CF	颯	98CF	颯	≡	r-simp		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
626C	扬	626C	扬	≡	r-simp		identity
626C	扬	63DA	揚	→	trad		
				←	simp		
626C	扬	656D	𪔐	↔	blocked		
626C	扬	98BA	颯	↔	blocked		
626C	扬	98CF	颯	↔	blocked		
63DA	揚	63DA	揚	≡	r-trad		identity
63DA	揚	656D	𪔐	↔	blocked		
63DA	揚	98BA	颯	↔	blocked		
63DA	揚	98CF	颯	↔	blocked		
656D	𪔐	656D	𪔐	≡	r-both		identity
656D	𪔐	98BA	颯	↔	blocked		
656D	𪔐	98CF	颯	↔	blocked		
98BA	颯	98BA	颯	≡	r-trad		identity
98BA	颯	98CF	颯	→	simp		
				←	trad		
98CF	颯	98CF	颯	≡	r-simp		identity

19. The code point U+64E1 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd). Note that U+64E1 is mostly used in a Korean context.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
62AC	抬	62AC	抬	≡	r-both		identity
62AC	抬	64E1	擡	→	blocked		
				←	simp		
64E1	擡	64E1	擡	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
62AC	抬	62AC	抬	≡	r-both		identity
62AC	抬	64E1	擡	↔	blocked		
64E1	擡	64E1	擡	≡	r-both		identity

The correlation between U+62AC and U+64E1 is confirmed by UniHan (kSemanticVariant), but the difference in mappings cannot be clarified.

20. The code point U+637F was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between dotAsia (isolated ‘r-both’) and CLGR7 (table follows). Note that U+637F is mostly used in a Korean context.

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
637F	棲	637F	棲	≡ r-neither		identity
637F	棲	6816	栖	→ simp ← blocked		
637F	棲	68F2	棲	→ trad ← blocked		
6816	栖	6816	栖	≡ r-both		identity
6816	栖	68F2	棲	→ trad ← simp		
68F2	棲	68F2	棲	≡ r-trad		identity

Unihan does not bring any clarification; this needs further study.

21. The code point U+6667 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
6667	皓	6667	皓	≡ r-neither		identity
6667	皓	66A0	暘	↔ blocked		
6667	皓	7693	皓	→ both		
				← blocked		
6667	皓	769C	皜	↔ blocked		
66A0	暘	66A0	暘	≡ r-trad		identity
66A0	暘	7693	皓	→ simp		
				← blocked		
66A0	暘	769C	皜	↔ blocked		
7693	皓	7693	皓	≡ r-both		identity
7693	皓	769C	皜	→ blocked		
				← simp		
769C	皜	769C	皜	≡ r-trad		identity
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
6667	皓	6667	皓	≡ r-both		identity
6667	皓	66A0	暘	↔ blocked		
6667	皓	7693	皓	↔ blocked		
6667	皓	769C	皜	↔ blocked		
66A0	暘	66A0	暘	≡ r-trad		identity

66A0	𩇛	7693	皓	→ simp		
				← blocked		
66A0	𩇛	769C	皓	↔ blocked		
7693	皓	7693	皓	≡ r-both		identity
7693	皓	769C	皓	→ blocked		
				← simp		
769C	皓	769C	皓	≡ r-trad		identity

Unihan does not bring any clarification; this needs further study.

22. The code points U+6900 and U+76CC were included in CLGR7 because of their IICORE property but are treated differently between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6900	椀	6900	椀	≡	r-simp		identity
6900	椀	76CC	盃	↔	blocked		
6900	椀	7897	碗	→	trad		
				←	blocked		
6900	椀	92FA	碗	↔	blocked		
76CC	盃	76CC	盃	≡	r-neither		identity
76CC	盃	7897	碗	→	both		
				←	blocked		
76CC	盃	92FA	碗	↔	blocked		
7897	碗	7897	碗	≡	r-both		identity
7897	碗	92FA	碗	→	blocked		
				←	simp		
92FA	碗	92FA	碗	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6900	椀	6900	椀	≡	r-both		identity
6900	椀	76CC	盃	↔	blocked		
6900	椀	7897	碗	↔	blocked		
6900	椀	92FA	碗	↔	blocked		
76CC	盃	76CC	盃	≡	r-both		identity

76CC	盃	7897	碗	↔	blocked		
76CC	盃	92FA	碗	↔	blocked		
7897	碗	7897	碗	≡	r-both		identity
7897	碗	92FA	碗	→	blocked		
				←	simp		
92FA	碗	92FA	碗	≡	r-trad		identity

While in Unihan there is a kSemanticVariant mapping between U+6900 and U+7897, the 'trad' mapping suggested by CLGR7 seems dubious, and the dotAsia mapping seems preferable. The other differences require more study.

23. The code point U+6A53 was included in CLGR7 because of its membership in IICORE but is treated differently between dotAsia and CLGR7 (table follows). In dotAsia, both are singleton reflexive variants of type ‘r-both’.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6A53	𣎵	6A53	𣎵	≡	r-neither		identity
6A53	𣎵	8563	𣎶	→	both		
				←	blocked		
8563	𣎶	8563	𣎶	≡	r-both		identity

UniHan has a kSemanticVariant mapping between U+6A53 and U+8563. However, the propose mapping requires further study.

24. The code point U+8B0C was included in CLGR7 but is treated differently between dotAsia and CLGR7 (table follows). In dotAsia, both code points are singleton reflexive variants of type ‘r-both’.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6B4C	歌	6B4C	歌	≡	r-both		identity
6B4C	歌	8B0C	𪛗	→	blocked		
				←	simp		
8B0C	𪛗	8B0C	𪛗	≡	r-trad		identity

UniHan has a KSemanticVariant mapping between U+6B4C and U+8B0C, but not traditional or simplified mapping is hinted, therefore this change requires further study.

25. The code point U+6EDD was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6CF7	泷	6CF7	泷	≡	r-simp		identity
6CF7	泷	6EDD	滝	→	blocked		
				←	simp		
6CF7	泷	7027	瀧	→	trad		
				←	simp		
6EDD	滝	6EDD	滝	≡	r-neither		identity
6EDD	滝	7027	瀧	→	trad		
				←	blocked		
7027	瀧	7027	瀧	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
6CF7	泷	6CF7	泷	≡	r-simp		identity
6CF7	泷	6EDD	滝	→	blocked		
				←	simp		
6CF7	泷	7027	瀧	→	trad		
				←	simp		
6EDD	滝	6EDD	滝	≡	r-trad		identity
6EDD	滝	7027	瀧	↔	blocked		
7027	瀧	7027	瀧	≡	r-trad		identity

There is no support in Unihan for a traditional mapping between U+6EDD and U+7027 as suggested by CLGR7. This requires further study.

26. The code point U+7AC3 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7076	灶	7076	灶	≡	r-both		identity
7076	灶	7AC3	竈	→	blocked		
				←	simp		
7076	灶	7AC8	竈	→	blocked		
				←	both		
7AC3	竈	7AC3	竈	≡	r-neither		identity
7AC3	竈	7AC8	竈	→	trad		
				←	blocked		
7AC8	竈	7AC8	竈	≡	r-neither		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7076	灶	7076	灶	≡	r-both		identity
7076	灶	7AC3	竈	→	blocked		
				←	simp		
7076	灶	7AC8	竈	→	blocked		
				←	both		
7AC3	竈	7AC3	竈	≡	r-trad		identity
7AC3	竈	7AC8	竈	↔	blocked		
7AC8	竈	7AC8	竈	≡	r-neither		identity

Unihan does not bring any clarification; this needs further study.

27. The code point U+932C was included in CLGR7 because of its membership in IICORE but is treated differently between dotAsia and CLGR7 (table follows). In dotAsia U+932C is a singleton reflexive variant of type 'r-both'.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
70BC	炼	70BC	炼	≡	r-simp		identity
70BC	炼	7149	煉	→	trad		
				←	simp		
70BC	炼	932C	鍊	→	blocked		
				←	simp		
70BC	炼	934A	鍊	→	blocked		
				←	simp		
7149	煉	7149	煉	≡	r-trad		identity
7149	煉	932C	鍊	→	blocked		
				←	trad		
7149	煉	934A	鍊	↔	blocked		
932C	鍊	932C	鍊	≡	r-neither		identity
932C	鍊	934A	鍊	↔	blocked		
934A	鍊	934A	鍊	≡	r-trad		identity

Unihan does not bring any clarification; this needs further study.

28. The code point U+7200 was included in CLGR7 because of its membership in IICORE but is treated differently between dotAsia and CLGR7 (table follows). In dotAsia both code points are singleton reflexive variants of type 'r-both'.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7200	𤇗	7200	𤇗	≡	r-neither		identity
7200	𤇗	8D6B	𤇗	→	both		
				←	blocked		
8D6B	𤇗	8D6B	𤇗	≡	r-both		identity

Unihan does not bring any clarification; this needs further study.

29. The code point U+734F was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
734F	𤇚	734F	𤇚	≡	r-neither		identity
734F	𤇚	8C98	𤇚	→	both		
				←	blocked		
8C98	𤇚	8C98	𤇚	≡	r-both		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
734F	𤇚	734F	𤇚	≡	r-trad		identity
734F	𤇚	8C98	𤇚	→	simp		
				←	blocked		
8C98	𤇚	8C98	𤇚	≡	r-both		identity

Unihan does not bring any clarification; this needs further study.

30. The code point U+73E1 was included in CLGR7 because of its membership in IICORE but is treated differently between dotAsia and CLGR7 (table follows). In dotAsia both code points are singleton reflexive variants of type 'r-both'.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
73E1	𪚳	73E1	𪚳	≡	r-neither		identity
73E1	𪚳	7434	𪚴	→	both		
				←	blocked		
7434	𪚴	7434	𪚴	≡	r-both		identity

Unihan does not bring any clarification; this needs further study.

31. The code point U+74C8 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
740D	琍	740D	琍	≡ r-trad		identity
740D	琍	7483	璃	→ simp ← blocked		
740D	琍	74C8	璿	→ blocked ← trad		
7483	璃	7483	璃	≡ r-both		identity
7483	璃	74C8	璿	→ blocked ← simp		
74C8	璿	74C8	璿	≡ r-neither		identity
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
740D	琍	740D	琍	≡ r-trad		identity
740D	琍	7483	璃	→ simp ← blocked		
740D	琍	74C8	璿	↔ blocked		
7483	璃	7483	璃	≡ r-both		identity
7483	璃	74C8	璿	↔ blocked		
74C8	璿	74C8	璿	≡ r-both		identity

Although Unihan describes a kSemanticVariant between U+74C8 and U+7483, there is no traditional or simplified mapping implied for those code points.

32. The code point U+757A was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
757A	疆	757A	疆	≡	r-neither		identity
757A	疆	7586	疆	→	both		
				←	blocked		
7586	疆	7586	疆	≡	r-both		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
757A	疆	757A	疆	≡	r-both		identity
757A	疆	7586	疆	↔	blocked		
7586	疆	7586	疆	≡	r-both		identity

Unihan has a KSemanticVariant between U+757A and U+7586. But as for many others, this does not determine whether U+757A should be 'pre-empted'.

33. The code point U+764E was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
75EB	痲	75EB	痲	≡	r-simp		identity
75EB	痲	7647	痲	→	trad		
				←	simp		
75EB	痲	764E	痲	→	blocked		
				←	simp		
7647	痲	7647	痲	≡	r-trad		identity
7647	痲	764E	痲	→	blocked		
				←	trad		
764E	痲	764E	痲	≡	r-neither		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
75EB	痲	75EB	痲	≡	r-simp		identity
75EB	痲	7647	痲	→	trad		
				←	simp		
75EB	痲	764E	痲	↔	blocked		
7647	痲	7647	痲	≡	r-trad		identity
7647	痲	764E	痲	↔	blocked		
764E	痲	764E	痲	≡	r-both		identity

Unihan does not bring any clarification; this needs further study.

34. The code point U+89A9 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7779	睹	7779	睹	≡	r-both		identity
7779	睹	89A9	覩	→	blocked		
				←	both		
89A9	覩	89A9	覩	≡	r-neither		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7779	睹	7779	睹	≡	r-both		identity
7779	睹	89A9	覩	↔	blocked		
89A9	覩	89A9	覩	≡	r-both		identity

Unihan has a KSemanticVariant between U+7779 and U+89A9. But as for many others, this does not determine whether U+89A9 should be ‘pre-empted’.

35. The code point U+7C83 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
79D5	𪗇	79D5	𪗇	≡	r-both		identity
79D5	𪗇	7C83	𪗈	→	blocked		
				←	both		
7C83	𪗈	7C83	𪗈	≡	r-neither		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
79D5	𪗇	79D5	𪗇	≡	r-both		identity
79D5	𪗇	7C83	𪗈	↔	blocked		
7C83	𪗈	7C83	𪗈	≡	r-both		identity

Unihan has a KSemanticVariant between U+79D5 and U+7C83. But as for many others, this does not determine whether U+7C83 should be ‘pre-empted’.

36. The code point U+7B6F was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7B6F	筋	7B6F	筋	≡	r-neither		identity
7B6F	筋	7BB8	箸	→	both		
				←	blocked		
7BB8	箸	7BB8	箸	≡	r-both		identity, reflexive
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7B6F	筋	7B6F	筋	≡	r-both		identity
7B6F	筋	7BB8	箸	↔	blocked		
7BB8	箸	7BB8	箸	≡	r-both		identity

Unihan has a KSemanticVariant between U+7B6F and U+7BB8. But as for many others, this does not determine whether U+7B6F should be ‘pre-empted’.

37. The code point U+7DDC was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7DBF	綿	7DBF	綿	≡	r-trad		identity
7DBF	綿	7DDC	縣	→	blocked		
				←	trad		
7DBF	綿	7EF5	绵	→	simp		
				←	trad		
7DDC	縣	7DDC	縣	≡	r-neither		identity
7DDC	縣	7EF5	绵	→	simp		
				←	blocked		
7EF5	绵	7EF5	绵	≡	r-simp		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
7DBF	綿	7DBF	綿	≡	r-trad		identity
7DBF	綿	7DDC	縣	↔	blocked		
7DBF	綿	7EF5	绵	→	simp		
				←	trad		
7DDC	縣	7DDC	縣	≡	r-both		identity
7DDC	縣	7EF5	绵	↔	blocked		
7EF5	绵	7EF5	绵	≡	r-simp		identity

Unihan has a KSemanticVariant between U+7DBF and U+7DDC. But as for many others, this does not determine whether U+7DDC should be 'pre-empted'.

38. The code point U+9771 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
8095	𠄎	8095	𠄎	≡	r-trad		identity
8095	𠄎	976D	𠄎	↔	blocked		
8095	𠄎	9771	𠄎	↔	blocked		
8095	𠄎	97CC	𠄎	↔	blocked		
8095	𠄎	97E7	𠄎	→	simp		
				←	blocked		
976D	𠄎	976D	𠄎	≡	r-neither		identity
976D	𠄎	9771	𠄎	↔	blocked		
976D	𠄎	97CC	𠄎	→	trad		
				←	blocked		
976D	𠄎	97E7	𠄎	→	simp		
				←	blocked		
9771	𠄎	9771	𠄎	≡	r-neither		identity
9771	𠄎	97CC	𠄎	→	trad		
				←	blocked		
9771	𠄎	97E7	𠄎	→	simp		
				←	blocked		
97CC	𠄎	97CC	𠄎	≡	r-trad		identity

97CC	韌	97E7	韌	→ simp ← trad		
97E7	韌	97E7	韌	≡ r-simp		identity
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
8095	𠄎	8095	𠄎	≡ r-trad		identity
8095	𠄎	976D	韌	↔ blocked		
8095	𠄎	9771	韌	↔ blocked		
8095	𠄎	97CC	韌	↔ blocked		
8095	𠄎	97E7	韌	→ simp ← blocked		
976D	韌	976D	韌	≡ r-neither		identity
976D	韌	9771	韌	→ blocked ← trad		
976D	韌	97CC	韌	→ trad ← blocked		
976D	韌	97E7	韌	→ simp ← blocked		
9771	韌	9771	韌	≡ r-neither		identity
9771	韌	97CC	韌	↔ blocked		
9771	韌	97E7	韌	→ simp		

				← blocked		
97CC	韌	97CC	韌	≡ r-trad		identity
97CC	韌	97E7	韌	→ simp		
				← trad		
97E7	韌	97E7	韌	≡ r-simp		identity

in Unihan, U+9771 is a zVariant of U+97CC. The variation in 'trad' mapping between the two LGRs for U+9771 cannot be determined. Further study is needed.

39. The code point U+994D was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
81B3	膳	81B3	膳	≡	r-both		identity
81B3	膳	994D	饊	→	blocked		
				←	both		
994D	饊	994D	饊	≡	r-neither		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
81B3	膳	81B3	膳	≡	r-both		identity
81B3	膳	994D	饊	↔	blocked		
994D	饊	994D	饊	≡	r-both		identity

Unihan has a KSemanticVariant between U+81B3 and U+994D. But as for many others, this does not determine whether U+994D should be ‘pre-empted’.

40. The code point U+8420 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between dotAsia (table follows) and CLGR7 (it is just 'r-both' with no additional mapping).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
840C	萌	840C	萌	≡	r-both		identity
840C	萌	8420	萌	→	blocked		
				←	simp		
840C	萌	8544	蒔	→	blocked		
				←	simp		
8420	萌	8420	萌	≡	r-trad		identity
8420	萌	8544	蒔	↔	blocked		
8544	蒔	8544	蒔	≡	r-trad		identity

In Unihan U+8420 is a zVariant of U+U+840C, therefore some variant mapping is expected (as done in the dotAsia table). The dotAsia table seems preferable.

41. The code point U+84DA was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
84DA	𑖓	84DA	𑖓	≡	r-neither		identity
84DA	𑖓	84E8	𑖔	→	both		
				←	blocked		
84E8	𑖔	84E8	𑖔	≡	r-both		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
84DA	𑖓	84DA	𑖓	≡	r-both		identity
84DA	𑖓	84E8	𑖔	↔	blocked		
84E8	𑖔	84E8	𑖔	≡	r-both		identity

Unihan does not bring any clarification; this needs further study.

42. The code point U+8EE2 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
8EE2	𠄎	8EE2	𠄎	≡ r-neither		identity
8EE2	𠄎	8F49	轉	→ trad ← blocked		
8EE2	𠄎	8F6C	转	→ simp ← blocked		
8F49	轉	8F49	轉	≡ r-trad		identity
8F49	轉	8F6C	转	→ simp ← trad		
8F6C	转	8F6C	转	≡ r-simp		identity
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
8EE2	𠄎	8EE2	𠄎	≡ r-trad		identity
8EE2	𠄎	8F49	轉	→ trad ← blocked		
8EE2	𠄎	8F6C	转	→ simp ← blocked		
8F49	轉	8F49	轉	≡ r-trad		identity
8F49	轉	8F6C	转	→ simp ← trad		
8F6C	转	8F6C	转	≡ r-simp		identity

In Unihan U+8EE2 is a zVariant of U+8F49, as such the mapping of the CLGR7 makes more sense.

43. The code point U+8FBA was included in CLGR7 because of its membership in IICORE but is treated differently between dotAsia (where it is a singleton reflexive variant of type 'r-both') and CLGR7 (table follows).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
8FB9	边	8FB9	边	≡ r-simp		identity
8FB9	边	8FBA	辺	→ blocked ← simp		
8FB9	边	9089	邊	→ blocked ← simp		
8FB9	边	908A	邊	→ trad ← simp		
8FBA	辺	8FBA	辺	≡ r-neither		identity
8FBA	辺	9089	邊	↔ blocked		
8FBA	辺	908A	邊	→ trad ← blocked		
9089	邊	9089	邊	≡ r-neither		identity
9089	邊	908A	邊	→ trad ← blocked		
908A	邊	908A	邊	≡ r-trad		identity

In Unihan U+8FBA is a zVariant of U+908A, therefore the CLGR7 seems to make more sense.

44. The code point U+98C7 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
98C6	𩇛	98C6	𩇛	≡	r-trad		identity
98C6	𩇛	98C7	𩇛	→	blocked		
				←	trad		
98C6	𩇛	98C8	𩇛	→	blocked		
				←	trad		
98C6	𩇛	98D9	𩇛	→	simp		
				←	trad		
98C6	𩇛	98DA	𩇛	→	blocked		
				←	trad		
98C7	𩇛	98C7	𩇛	≡	r-neither		identity
98C7	𩇛	98C8	𩇛	↔	blocked		
98C7	𩇛	98D9	𩇛	→	simp		
				←	blocked		
98C7	𩇛	98DA	𩇛	↔	blocked		
98C8	𩇛	98C8	𩇛	≡	r-neither		identity
98C8	𩇛	98D9	𩇛	↔	blocked		
98C8	𩇛	98DA	𩇛	→	simp		
				←	blocked		

98D9	飊	98D9	飊	≡	r-simp		identity
98D9	飊	98DA	飗	↔	blocked		
98DA	飗	98DA	飗	≡	r-simp		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
98C6	飊	98C6	飊	≡	r-trad		identity
98C6	飊	98C7	飗	→	blocked		
				←	trad		
98C6	飊	98C8	飗	→	blocked		
				←	trad		
98C6	飊	98D9	飊	→	simp		
				←	trad		
98C6	飊	98DA	飗	→	blocked		
				←	trad		
98C7	飗	98C7	飗	≡	r-neither		identity
98C7	飗	98C8	飗	↔	blocked		
98C7	飗	98D9	飊	↔	blocked		
98C7	飗	98DA	飗	→	simp		
				←	blocked		
98C8	飗	98C8	飗	≡	r-neither		identity
98C8	飗	98D9	飊	↔	blocked		

98C8	風	98DA	風	→ simp ← blocked		
98D9	飊	98D9	飊	≡ r-simp		identity
98D9	飊	98DA	風	↔ blocked		
98DA	風	98DA	風	≡ r-simp		identity

Unihan does not bring any clarification; this needs further study.

45. The code point U+99C5 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
99C5	馱	99C5	馱	≡ r-neither		identity
99C5	馱	9A5B	驛	→ trad ← blocked		
99C5	馱	9A7F	驛	→ simp ← blocked		
9A5B	驛	9A5B	驛	≡ r-trad		identity
9A5B	驛	9A7F	驛	→ simp ← trad		
9A7F	驛	9A7F	驛	≡ r-simp		identity
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
99C5	馱	99C5	馱	≡ r-trad		identity
99C5	馱	9A5B	驛	→ trad ← blocked		
99C5	馱	9A7F	驛	→ simp ← blocked		
9A5B	驛	9A5B	驛	≡ r-trad		identity
9A5B	驛	9A7F	驛	→ simp ← trad		
9A7F	驛	9A7F	驛	≡ r-simp		identity

In Unihan, U+99C5 is a zVariant of U+9A5B. As such the CLGR7 mapping makes more sense.

46. The code point U+9D44 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
9D1F	鴉	9D1F	鴉	≡ r-trad		identity
9D1F	鴉	9D44	鴉	→ blocked ← trad		
9D1F	鴉	9E31	鴉	→ simp ← trad		
9D44	鴉	9D44	鴉	≡ r-neither		identity
9D44	鴉	9E31	鴉	→ simp ← blocked		
9E31	鴉	9E31	鴉	≡ r-simp		identity
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
9D1F	鴉	9D1F	鴉	≡ r-trad		identity
9D1F	鴉	9D44	鴉	↔ blocked		
9D1F	鴉	9E31	鴉	→ simp ← trad		
9D44	鴉	9D44	鴉	≡ r-both		identity
9D44	鴉	9E31	鴉	↔ blocked		
9E31	鴉	9E31	鴉	≡ r-simp		identity

Unihan indicates that U+9D44 is kDefinition variant of U+9D1F, but this does not provide a solution to determine which of the two LGRs is better.

47. The code point U+9F62 was included in CLGR7 because of its membership in IICORE but has been assigned different types for its variant mappings between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
9F61	齒令	9F61	齒令	≡ r-trad		identity, reflexive
9F61	齒令	9F62	齒令	→ blocked ← trad		
9F61	齒令	9F84	齒令	→ simp ← trad		
9F62	齒令	9F62	齒令	≡ r-neither		identity
9F62	齒令	9F84	齒令	→ simp ← blocked		
9F84	齒令	9F84	齒令	≡ r-simp		identity, reflexive
Source	Glyph	Target	Glyph	Type(s)	Ref	Comment
9F61	齒令	9F61	齒令	≡ r-trad		identity
9F61	齒令	9F62	齒令	→ blocked ← trad		
9F61	齒令	9F84	齒令	→ simp ← trad		
9F62	齒令	9F62	齒令	≡ r-trad		identity
9F62	齒令	9F84	齒令	→ simp ← blocked		
9F84	齒令	9F84	齒令	≡ r-simp		identity

In Unihan, U+9FC2 is a zVariant of U+9F61. As such the CLGR7 mapping makes more sense.

5.3 Variant sets with pre-existing code points but different variant types

For the following variant sets, all code points were already present in the original .cn/tw tables and were not the result of augmentations to sync with the dotAsia repertoire. However, the assigned variant types do not agree. In other words, these variant sets show deviation between CLGR7 and the original sources.

The IP has not yet evaluated the merits of the changes, and before doing so, would like to request the CGP to provide some background information, rationale, references etc. that explain these differences.

1. These 2 code points: U+5B0E and U+5B14 form their own variant set in dotAsia.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5B0E	媿	5B0E	媿	≡	r-both		identity
5B0E	媿	5B14	媿	↔	blocked		
5B14	媿	5B14	媿	≡	r-both		identity

Unihan does not provide any clues concerning this. Further study is needed.

2. In the CLGR7 (1st), U+752F is added in the following set with mapping of type 'blocked' type (except 'r-both' where mapped to itself). This was not the case in .cn. In dotAsia U+752F is a singleton reflexive mapping of type 'r-both'. In addition, the types of the mappings between U+5BD7-U+5BE7 and U+5B81-U+5BD7 are also different from dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5B81	宁	5B81	宁	≡	r-both		identity
5B81	宁	5BCD	寗	→	blocked		
				←	simp		
5B81	宁	5BD5	寧	→	blocked		
				←	simp		

5B81	宁	5BD7	寧	↔	blocked		
5B81	宁	5BDC	寧	→	blocked		
				←	simp		
5B81	宁	5BE7	寧	→	trad		
				←	simp		
5B81	宁	752F	甯	↔	blocked		
5BCD	寧	5BCD	寧	≡	r-trad		identity
5BCD	寧	5BD5	寧	↔	blocked		
5BCD	寧	5BD7	寧	↔	blocked		
5BCD	寧	5BDC	寧	↔	blocked		
5BCD	寧	5BE7	寧	↔	blocked		
5BCD	寧	752F	甯	↔	blocked		
5BD5	寧	5BD5	寧	≡	r-neither		identity
5BD5	寧	5BD7	寧	↔	blocked		
5BD5	寧	5BDC	寧	↔	blocked		
5BD5	寧	5BE7	寧	→	trad		
				←	blocked		
5BD5	寧	752F	甯	↔	blocked		
5BD7	寧	5BD7	寧	≡	r-both		identity

5BD7	寧	5BDC	寧	↔	blocked		
5BD7	寧	5BE7	寧	↔	blocked		
5BD7	寧	752F	甯	↔	blocked		
5BDC	寧	5BDC	寧	≡	r-neither		identity
5BDC	寧	5BE7	寧	→	trad		
				←	blocked		
5BDC	寧	752F	甯	↔	blocked		
5BE7	寧	5BE7	寧	≡	r-trad		identity
5BE7	寧	752F	甯	↔	blocked		
752F	甯	752F	甯	≡	r-both		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
5B81	宁	5B81	宁	≡	r-both		identity
5B81	宁	5BCD	寧	→	blocked		
				←	simp		
5B81	宁	5BD5	寧	→	blocked		
				←	simp		
5B81	宁	5BD7	寧	→	blocked		
				←	simp		
5B81	宁	5BDC	寧	→	blocked		
				←	simp		
5B81	宁	5BE7	寧	→	trad		

				← simp		
5BCD	寧	5BCD	寧	≡ r-trad		identity
5BCD	寧	5BD5	寧	↔ blocked		
5BCD	寧	5BD7	寧	↔ blocked		
5BCD	寧	5BDC	寧	↔ blocked		
5BCD	寧	5BE7	寧	↔ blocked		
5BD5	寧	5BD5	寧	≡ r-neither		identity
5BD5	寧	5BD7	寧	↔ blocked		
5BD5	寧	5BDC	寧	↔ blocked		
5BD5	寧	5BE7	寧	→ trad		
				← blocked		
5BD7	寧	5BD7	寧	≡ r-both		identity
5BD7	寧	5BDC	寧	↔ blocked		
5BD7	寧	5BE7	寧	→ trad		
				← blocked		
5BDC	寧	5BDC	寧	≡ r-neither		identity
5BDC	寧	5BE7	寧	→ trad		
				← blocked		
5BE7	寧	5BE7	寧	≡ r-trad		identity

Unihan supports the addition of U+752F to the set. There is however no evidence concerning the mapping changes concerning the pairs (U+5B81, U+5BD7) and (U+5BD7, 5BE7). Further study is needed.

3. In these tables, the types of the mappings between U+67A3 and U+6806 or U+68D7 are reversed between CLGR7 (1st) and dotAsia (2nd).

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
67A3	𣎵	67A3	𣎵	≡	r-simp		identity
67A3	𣎵	6806	𣎵	→	blocked		
				←	simp		
67A3	𣎵	68D7	𣎶	→	trad		
				←	simp		
6806	𣎵	6806	𣎵	≡	r-trad		identity
6806	𣎵	68D7	𣎶	↔	blocked		
68D7	𣎶	68D7	𣎶	≡	r-trad		identity
Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
67A3	𣎵	67A3	𣎵	≡	r-simp		identity
67A3	𣎵	6806	𣎵	→	trad		
				←	simp		
67A3	𣎵	68D7	𣎶	→	blocked		
				←	simp		
6806	𣎵	6806	𣎵	≡	r-trad		identity
6806	𣎵	68D7	𣎶	↔	blocked		

68D7	棗	68D7	棗	≡	r-trad		identity
------	---	------	---	---	--------	--	----------

Unihan confirms the CLGR7 option (U+68D7 kTraditionalVariant for U+67A3).

- In the CLGR7 (table follows), U+68C5 maps to U+67C4 with a type of 'both'. This is different from both .cn and dotAsia where it maps to itself as 'r-both' type and where there is no variant relation between these code points.

Source	Glyph	Target	Glyph		Type(s)	Ref	Comment
67C4	柄	67C4	柄	≡	r-both		identity
67C4	柄	68C5	棟	→	blocked		
				←	both		
68C5	棟	68C5	棟	≡	r-neither		identity

Unihan defines kSemanticVariant mapping between U+67C4 and U+68C5 which would support the CLGR7 mapping.

6 Variant Mappings that may overproduce allocatable labels

In the Chinese LGR, the variant mappings and WLE rules are designed with the assumption that given any valid input label, there would be at most three resulting allocatable labels -- the original label, an all-simplified label, and an all-traditional label. This is achieved using variant mappings of having at most one instance in each of the following set of types:

- trad, r-trad, both, r-both
- simp, r-simp, both, r-both

However, in CLGR7, there are 196 code points (attached file CLGR-Overproducing-Variants-20160530.txt) with variant types that violate the above constraint. This would lead to overproduction of variant labels with an "allocatable" status.

An example would be:

```
<char cp="53F0" tag="sc:Hani" ref="0 100 101 102 103 104" >
  <var cp="53F0" type="r-both" comment="identity,reflexive" />
  <var cp="6AAF" type="trad" />
  <var cp="7C49" type="block" />
  <var cp="81FA" type="trad" />
  <var cp="98B1" type="trad" />
```

</char>

Using the one-line notation in the attached file, the above is represented as:

53F0[台] trad=> 6AAF[檯] trad=> 81FA[臺] r-both=> 53F0[台] trad=> 98B1[颱]

An input label of 台湾 (53F0 6E7E) would result in 5 allocatable variant labels (action numbers indexed per sequence order of the <action> elements in the XML file: 0 to 5):

- **Variant: (檯灣) (6AAF 7063): [trad] ==> allocatable due to Action[2]**
- Variant: (檯灣) (6AAF 6E7E): [trad r-simp] ==> blocked due to Action[4]
- **Variant: (臺灣) (81FA 7063): [trad] ==> allocatable due to Action[2]**
- Variant: (臺灣) (81FA 6E7E): [trad r-simp] ==> blocked due to Action[4]
- **Variant: (台灣) (53F0 7063): [trad r-both] ==> allocatable due to Action[2]**
- **Variant: (台湾) (53F0 6E7E): [r-both r-simp] ==> allocatable due to Action[1]**
- **Variant: (颱灣) (98B1 7063): [trad] ==> allocatable due to Action[2]**
- Variant: (颱灣) (98B1 6E7E): [trad r-simp] ==> blocked due to Action[4]
- Variant: (臺灣) (7C49 7063): [trad block] ==> blocked due to Action[0]
- Variant: (臺灣) (7C49 6E7E): [r-simp block] ==> blocked due to Action[0]

Some of the above "allocatable" labels are unnecessary from a semantic standpoint.

In at least some of these 196 code points, the reason for the variant type assignments appears to be due to a simplified code point having multiple traditional variants. If so, it may be an acceptable trade-off to eliminate the multiple traditional mappings, and let registrants who need a specific traditional variant label apply for the specific traditional label.

If that was the argument, most of these cases (except perhaps for a few cases such as "Taiwan") can be fixed by not having multiple traditional mappings. Registrants who want a specific traditional label should apply for the traditional string, which should give the right simplified string, and won't over-generate.

An example would be a label involving two of the code points that exhibit this issue:

66F2[曲] r-both=> 66F2[曲] trad=> 9EB4[麩]

9709[霉] r-both=> 9709[霉] trad=> 9EF4[黴]

The label 红曲霉 (7EA2 66F2 9709) "red yeast" would yield 5 allocatable labels:

- Variant: (红麩霉) (7EA2 9EAF 9709): [r-both block r-simp] ==> blocked due to Action[0]
- Variant: (红麩黴) (7EA2 9EAF 9EF4): [trad block r-simp] ==> blocked due to Action[0]
- **Variant: (红曲霉) (7EA2 66F2 9709): [r-both r-simp] ==> allocatable due to Action[1]**
- Variant: (红曲黴) (7EA2 66F2 9EF4): [trad r-both r-simp] ==> blocked due to Action[4]
- Variant: (红麩霉) (7EA2 9EB9 9709): [r-both block r-simp] ==> blocked due to Action[0]
- Variant: (红麩黴) (7EA2 9EB9 9EF4): [trad block r-simp] ==> blocked due to Action[0]

- Variant: (紅麴霉) (7EA2 9EB4 9709): [trad r-both r-simp] ==> blocked due to Action[4]
- Variant: (紅麴黴) (7EA2 9EB4 9EF4): [trad r-simp] ==> blocked due to Action[4]
- Variant: (紅麴霉) (7D05 9EAF 9709): [trad r-both block] ==> blocked due to Action[0]
- Variant: (紅麴黴) (7D05 9EAF 9EF4): [trad block] ==> blocked due to Action[0]
- **Variant: (紅曲霉) (7D05 66F2 9709): [trad r-both] ==> allocatable due to Action[2]**
- **Variant: (紅曲黴) (7D05 66F2 9EF4): [trad r-both] ==> allocatable due to Action[2]**
- Variant: (紅麴霉) (7D05 9EB9 9709): [trad r-both block] ==> blocked due to Action[0]
- Variant: (紅麴黴) (7D05 9EB9 9EF4): [trad block] ==> blocked due to Action[0]
- **Variant: (紅麴霉) (7D05 9EB4 9709): [trad r-both] ==> allocatable due to Action[2]**
- **Variant: (紅麴黴) (7D05 9EB4 9EF4): [trad] ==> allocatable due to Action[2]**

Instead, if the variant types could be amended to the following:

66F2[曲] r-both=> 66F2[曲] **blocked=> 9EB4[麴]**

9709[霉] r-both=> 9709[霉] **blocked=> 9EF4[黴]**

The same input label 紅曲霉 (7EA2 66F2 9709) "red yeast" would yield 2 allocatable labels (omitting output labels that have been assigned a "blocked" disposition):

- **Variant: (紅曲霉) (7EA2 66F2 9709): [r-both r-simp] ==> allocatable due to Action[1]**
- **Variant: (紅曲霉) (7D05 66F2 9709): [trad r-both] ==> allocatable due to Action[2]**

This may not be desirable because 紅麴霉 (7D05 9EB4 9709) is perhaps more appropriate. In that case, it can be the applied-for label, which would then yield the following 2 allocatable labels:

- **Variant: (紅曲霉) (7EA2 66F2 9709): [simp r-both] ==> allocatable due to Action[1]**
- **Variant: (紅麴霉) (7D05 9EB4 9709): [r-trad r-both] ==> allocatable due to Action[2]**

Conclusion: Under the conservatism principle, LGRs should strive to minimize allocatable variants. The IP would like to urge the CGP to change the variant types of the affected code points to mitigate this issue, and/or provide strong evidence for the need of including exceptional cases with multiple allocatable variants.
