Considerations concerning the Chinese Root LGR

**Last updated: January 30, 2018**

Contents

[1 Summary 2](#_Toc505113484)

[2 Definitions 4](#_Toc505113485)

[2.1 CLGR10 4](#_Toc505113486)

[2.2 IICORE collection 4](#_Toc505113487)

[2.3 MSR-2 and MSR-3 CJK repertoire 4](#_Toc505113488)

[2.4 dotAsia LGR 4](#_Toc505113489)

[2.5 Unihan 5](#_Toc505113490)

[3 Repertoire considerations 5](#_Toc505113491)

[3.1 HKSCS additions – 2 code points 5](#_Toc505113492)

[3.2 TGSCC additions – 18 code points 6](#_Toc505113493)

[3.3 IICORE J and K Source characters with variants – 42 code points 7](#_Toc505113494)

[4 Variant considerations 8](#_Toc505113495)

[4.1 General 8](#_Toc505113496)

[4.2 Types of differences 10](#_Toc505113497)

[4.3 Notation and explanation used in the description of the variant set differences 10](#_Toc505113498)

[4.3.1. Table format for variant sets 10](#_Toc505113499)

[4.3.2. Source information 11](#_Toc505113500)

[5 Variant set differences by type of difference 11](#_Toc505113501)

[5.1 Additional repertoire (CLGR10 adds 62 code points not in dotAsia) 11](#_Toc505113502)

[5.1.1. Review of variant sets for out-of-repertoire variant beyond dotAsia – 42 sets 12](#_Toc505113503)

[5.1.2. Review of variant sets for additional HKSCS characters beyond dotAsia – 2 sets 13](#_Toc505113504)

[5.1.3. Review of variant sets for TGSCC characters beyond dotAsia – 18 sets 13](#_Toc505113505)

[5.1.4. Detailed review of variant sets modified by additions beyond dotAsia – 8 sets 14](#_Toc505113506)

[5.2 Code points added to CLGR10 through dotAsia but treated differently 26](#_Toc505113507)

[5.2.1. Summary of the list of additions treated differently – 45 items 26](#_Toc505113508)

[5.2.2. Detailed list of differences for dotAsia additions – 31 sets 28](#_Toc505113509)

[5.3 Variants sets with pre-existing code points pre-integrated with KLGR 61](#_Toc505113510)

[5.3.1. Table of differences from pre-integration – 146 groups of variant sets 61](#_Toc505113511)

[5.3.2. Detailed list of KLGR pre-integration discrepancies – 6 variant sets 66](#_Toc505113512)

[5.4 Variant sets with pre-existing code points but different variant types – 3 sets 72](#_Toc505113513)

[6 Mechanism for reducing multiple allocatable labels 74](#_Toc505113514)

# Summary

This document presents Integration Panel recommendations for the Chinese Root Zone LGR proposal together with an analysis of 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 (CLGR10) is represented by the following files:

* CGP-LGR-Proposal-1.1-201712.docx, [Proposal]
* Appendix A CGP Repertoire 201712.xlsx,
* Appendix B JGP Repertoire 201703.xlsx,
* Appendix C KGP Hanja Repertoire 201703.xlsx,
* Appendix D CGP Variant Mappings 201712.xlsx,
* Appendix E CGP Variant Review on new added Characters 2015~2016.xlsx,
* Proposed-LGR-Hani-1.1-20171212.xml

In evaluating this proposal, the IP performed a comparison to the dotAsia ZH set (see 2.4), which, like the CLGR10, attempts to cover both simplified and traditional Chinese labels. The IP also took into consideration the process of the Chinese Generation Panel (CGP) and the Korean General Panel (KGP) to synchronize their variant sets (also named as ‘pre-integration’).

The following list summarizes the analysis:

* The repertoire contains 19,745 code points, very close to the size of MSR-2 Hanzi set (19,850). It is a full superset of the dotAsia set, itself made of 19,683 code points.
* Of the 62 characters added to CLGR10 beyond dotAsia, 42 are added solely to complete variant sets and are correctly defined as “out-of-repertoire” variants, 18 are added as part of the Table of General Standard Chinese Character (TGSCC).
* The repertoire includes 2 characters not part of MSR-2; they will be included in MSR-3. (Note MSR-3 will add an additional Han character beyond these two).
* As identified in a previous version of this feedback, there is no sufficient explanation provided for the differences in variant sets between this LGR and, for example, existing second level LGRs. Many of these differences are still present and therefore should be carefully documented with a rationale supplied.
* A subset of these differences are due to the pre-integration with KLGR. See separate discussion.
* After a careful review of CLGR10 including associated files, the IP has identified the remaining problematic cases. This set is smaller than for previous versions of CLGR, so while the work is not complete, progress has been made.
* In order to establish that there is a rational and linguistic basis for assigning variants, any deviations from existing practice should be justified. If it is felt necessary or useful, some face-to-face or virtual meetings may be setup between IP and CGP to make progress on this issue.
* While the origin of some variants can be traced to the .asia, .cn, and, .tw sets (henceforth “dotAsia”, “dotcn”, and “dottw”), insufficient references or source information are provided for the modified or new variant sets. (The source for all variants, whether retained from second level or modified, should be unambiguously documented – either on a per-variant level, or globally, with any exceptions clearly marked).
* Because this LGR must be integrated with the rest of the CJK LGR (i.e. the Japanese and Korean LGR), the IP notes the progress made in aligning the variant sets between CLGR and KLGR; however, the documentation of the result of this process could be improved.
* Finally, the proposed mechanism for reducing allocatable labels should be fleshed out, with variant type assignments reflected in a separate copy of the XML. See Section 6.

**Integration Panel recommendations:**

* Further clarify the need to include all 18 characters of the TGSCC set by showing usage in Chinese context.
* Complete documentation for the origin of the proposed variant mappings in CLGR10, particularly where they differ from established second-level practice.
* Review variant sets that differ from second-level practice and provide a rationale for any differences, particularly in cases not forced by alignment with KLGR.
* Address issues noted in documenting the alignment with KLGR.
* Implement the proposed variant types for reducing overproduction of allocatable variants. If possible, please continue to maintain a separate parallel copy of the XML in order to isolate the variant type changes from any other change.
* Document the specific requirements behind any decision to retain multiple “trad” variants.
* 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. [[1]](#footnote-1)
* 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

# Definitions

## CLGR10

The term (CLGR10) represents the Proposed Chinese root LGR under review here, both in terms of repertoire and variant sets defined in the XML file. The term CLGR9 may be used to represent the earlier version of the Chinese LGR.

## 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.

## MSR-2 and MSR-3 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).

A pending update to the MSR will add three additional Han characters, code points: U+3A5C, U+58B5, and U+20B9F. The two first characters are part of CLGR10.

## dotAsia LGR

The dotAsia (ZH) IDN table is available at <https://www.iana.org/domains/idn-tables/tables/asia_zh_1.1.txt>. To compare this to CLGR10, the IP used a transcript of this table into the XML-format available at <https://www.icann.org/sites/default/files/packages/lgr/lgr-second-level-chinese-30aug16-en.xml> . This transcription was created as part of reference for 2nd level domain. It shares many features with the proposed root Chinese LGR. Technically, this transcription of the dotAsia (ZH) domain contains one additional CJK ideograph beyond the original dotAsia table: the code point U+9DC0 was added to complete a variant set. The original IDN table contains 19 683 Han ideographs. In the following comparisons of CLGR10 and dotAsia repertoire, the original 19 683 code points defined in dotAsia are used.

The dotAsia table (or its XML transcription) represents a useful reference for the comparison in terms of both the repertoire and the variant sets. It contains 19 683 (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.

The dotAsia repertoire is fully included in CLGR10. CLGR10 contains an additional 62 code points not included in dotAsia.

## 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 a very useful tool for validating the CLGR10 content.

# Repertoire considerations

The [Proposal] in section 5 describes the development of the repertoire, including detailing the various iterations from the initial repertoire to the current repertoire. The current repertoire can also be simply derived from the dotAsia repertoire by adding the following list of 62 characters:

* 2 HKSCS characters that were left out when processing HKIRC request
* 18 characters from the Table of General Standard Chinese Characters (TGSCC formerly known as Normalized Hanzi List for Common Use (NHCU) in previous drafts of CLGR)
* 42 IICORE characters from J source and K sources part of Chinese based variant sets. These are also part of draft versions of the JGP and KGP repertoires.

The 2 HKSCS characters (U+3A5C and U+58B5 墵), were not part of MSR-2; the pending new version of the MSR, MSR-3 will contain them. Both code points are members of CLGR10 variant sets.

(Note: MSR-3 will also add an additional Han character beyond these two: U+20B9F)



**Integration Panel recommendation:**

The GP is encouraged to review whether it is used in Chinese and / or has variant relations to code points in the CLGR10 repertoire). Note that it has a H source.

## HKSCS additions – 2 code points

The 2 characters from HKSCS can be described as follows: IRG Sources (GE sources excluded), IICORE value, and Unihan variants from Unihan, CLGR10 variants from [CLGR10]. 17 code points (out of 18) are members of CLGR10 variant sets. The yellow highlighting show variants introduced by CLGR10 in addition to Unihan. A code point marked in red indicates a Unihan variant not included in CLGR10.

| **No** | **UCS** | **Glyph** | **IRG Sources** | **IICORE** | **Unihan variants** | **CLGR10 variants** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 3A5C | 㩜 | G5,T3,JA,H |  | 3A2B, 64E5, 652C | 39DB, 63FD, 64E5, 652C |
| 2 | 58B5 | 墵 | T3,H |  | \* | 575B, 57EE, 58B0, 58C7, 58DC, 7F48, 7F4E |

\* U+58B5 has no variant per Unihan, but Unihan has a variant set (U+5081, U+575B, U+58B0, U+58C7, U+58DC, and U+7F48) having many common members with the CLGR10 variant set associated with U+58B5.

## TGSCC additions – 18 code points

The 18 characters from TGSCC can be described as follows: class values from section 5.2.3.1 of the [Proposal], IRG Sources (GE sources excluded), IICORE value, and Unihan variants from Unihan, CLGR10 variants from [CLGR10]. 17 code points (out of 18) are members of CLGR10 variant sets. The yellow highlighting show variants introduced by CLGR10 in addition to Unihan.

| **No** | **UCS** | **Glyph** | **Class** | **IRG Sources** | **IICORE** | **Unihan variants** | **CLGR10 variants** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 48BC | 䢼 | N | G3,T4 |  |  |  |
| 2 | 732F | 猯 | N | G5,T3,J0,K2 |  |  | 8C92 |
| 3 | 9EB9 | 麹 | N | J0 | AJ |  | 9EB4 |
| 4 | 5227 | 刧 | V | T3,J0,K1,H |  | 5226, 52AB | 5226, 523C, 52AB |
| 5 | 524F | 剏 | V | T3,J0,K2,H |  | 5231 | 521B, 5231, 5259, 5275 |
| 6 | 6060 | 恠 | V | T3,J0,K1 |  | 602A | 602A |
| 7 | 74A2 | 璢 | V | T3,J0,K2 |  | 7409 | 7409, 7460 |
| 8 | 750E | 甎 | V | T3,J0,K1,H |  | 78DA | 7816, 78DA |
| 9 | 754A | 畊 | V | T3,J0,K1,H |  | 8015 | 8015 |
| 10 | 7ADA | 竚 | V | T3,J0,K2 |  | 4F47 | 4F2B, 4F47 |
| 11 | 8262 | 艢 | V | T3,J0,K2,HB2 |  | 6AA3 | 6A2F, 6AA3 |
| 12 | 88B5 | 袵 | V | T3,J0,K2,H |  | 887D | 887D |
| 13 | 894D | 襍 | V | T3,J0,K1,H |  | 96DC | 6742, 96D1, 96DC, 96E5 |
| 14 | 8B0C | 謌 | V | T3,J0,K1,H |  | 6B4C | 6B4C |
| 15 | 8F19 | 輙 | V | T3,J0,K2,H |  | 8F12 | 8F12, 8F84 |
| 16 | 945A | 鑚 | V | T3,J0,K2 | CJ | 947D | 9246, 9409, 947D, 94BB |
| 17 | 984B | 顋 | V | T3,J0.K1,H |  | 816E | 816E |
| 18 | 9DC0 | 鷀 | V | G1,T4,J3,K2,H |  |  | 9DBF, 9E5A |

Findings concerning TGSCC additions:

* 16 characters have J0 sources (core Japanese Kanji set), among these 16 J0 sources 14 have the variant class (V) according to section 3.2.18 of the [Proposal]. Another character (code point U+9DC0) is part of a variant set (9DBF 鶿, 9DC0 鷀, 9E5A 鹚), although it is not recorded in Unihan (Unihan only recognizes a variant relationship between U+9DBF and U+9E5A).
* All Unihan variants are included in CLGR10, but the latter add many members.
* 15 characters have a TGSCC class ‘V’ and the last 3 have a class ‘N’ according to section 5.2.3.1 of the [Proposal].
* Many code points have an IRG source ‘H’ value, which may indicate use in Hong Kong or Guangzhou area (Cantonese dialect), this could be confirmed by CGP.

It would be useful to know the rationale for inclusion of these 18 characters because at first approach most of the characters are mostly used in Japanese context. In addition, a clarification on these abbreviations ‘N’ and ‘V’ would be desirable.

**Integration Panel recommendations:**

Further clarify the need to include all 18 characters of the TGSCC set by showing usage in Chinese context; specify the meaning of ‘N’ and ‘V’ in the table describing them in section 5.2.3.1.

## IICORE J and K Source characters with variants – 42 code points

The last 42 characters are IICORE characters from J source and K sources part of Chinese based variant sets. They are identified in section 5.2.3.2 of the [Proposal]. The 42 characters can be described as follows: IRG Sources (GE sources excluded), IICORE value, and Unihan variants from Unihan, CLGR10 variants from [CLGR10]. The yellow highlighting show variants introduced by CLGR10 in addition to Unihan. A code point marked in red indicates a Unihan variant not included in CLGR10.

| **No** | **UCS** | **Glyph** | **IRG Sources** | **IICORE** | **Unihan variant** | **CLGR10 variants** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 3960 | 㥠 | G5,T3,JA,K3 | CK | 8ADD | 8ADD, 8C1E |
| 2 | 4FAD | 侭 | J0 | AJ | 5118 | 5118, 5C3D, 76E1 |
| 3 | 51E6 | 処 | J0,K2 | AJ | 458F, 8655 | 5904, 8655 |
| 4 | 56A2 | 嚢 | J0 | AJ | 56CA | 56CA |
| 5 | 61F4 | 懴 | J0,K2 | CJ | 61FA | 5FCF, 61FA |
| 6 | 6442 | 摂 | J0 | AJ | 651D | 6315, 6444, 651D |
| 7 | 663B | 昻 | T3,K0 | AKP | 6602 | 6602 |
| 8 | 685C | 桜 | J0,K2 | AJ | 6AFB | 6A31, 6AFB |
| 9 | 685F | 桟 | J0 | AJ | 68E7 | 6808, 68E7, 8F4F |
| 10 | 6D9C | 涜 | J0,K2 | AJ | 7006 | 6E0E, 7006 |
| 11 | 6E8C | 溌 | J0,K2 | AJ | 6F51 | 6CFC, 6F51 |
| 12 | 731F | 猟 | J0,K2 | AJ | 7375 | 730E, 7375 |
| 13 | 784F | 硏 | T3,J3,K0 | AKP |  | 63C5, 7814 |
| 14 | 7A36 | 稶 | T4,K0 | AKP |  | 7A22 |
| 15 | 7B86 | 箆 | T3,J0,K2 | AJ |  | 7BE6 |
| 16 | 7C14 | 簔 | TF,J0 | CJ | 7C11 | 7C11, 84D1 |
| 17 | 7D9A | 続 | J0,K2 | AJ | 7E8C | 7E8C, 7EED |
| 18 | 7E4A | 繊 | J0,K2 | ATJ | 7E96 | 5B45, 7E34, 7E8E, 7E96, 7EA4 |
| 19 | 7E4B | 繋 | J0,K2 | AJ | 7E6B | 7E6B |
| 20 | 8133 | 脳 | J0 | AJ | 8166 | 8111, 8166 |
| 21 | 81D3 | 臓 | J0 | AJ | 81DF | 810F, 81DF, 9AD2 |
| 22 | 8217 | 舗 | J0 | AJ | 92EA | 8216, 92EA, 94FA |
| 23 | 839F | 莟 | G3,T3,J0,K1 | CJ |  | 83E1 |
| 24 | 83B5 | 莵 | J0 | CJ | 83DF | 83DF |
| 25 | 86CD | 蛍 | J0 | AJ | 87A2 | 8424, 87A2 |
| 26 | 8E99 | 躙 | G5,T3,J0,K1 | CJ |  | 8E8F, 8EAA |
| 27 | 9039 | 逹 | J0 | CJ | 9054 | 8FBE, 8FD6, 9054 |
| 28 | 91A4 | 醤 | J0 | AJ | 91AC | 9171, 91AC |
| 29 | 91C8 | 釈 | J0,K2 | AJ | 91CB | 91CA, 91CB |
| 30 | 9421 | 鐡 | J0,K2 | CJ | 9435 | 9244, 9295, 9435, 94C1 |
| 31 | 967A | 険 | J0,K2 | AJ | 96AA | 7877, 78B1, 7906, 9669, 96AA, 9E7C |
| 32 | 96B2 | 隲 | T3,J0 | CJ | 9A2D | 9A2D, 9A98 |
| 33 | 982C | 頬 | J0 | AJ | 9830 | 9830, 988A |
| 34 | 98EE | 飮 | J0,K0 | AKP | 98F2 | 98F2, 996E |
| 35 | 9A12 | 騒 | J0 | AJ | 9A37 | 9A37, 9A9A |
| 36 | 9A13 | 験 | J0,K2 | AJ | 9A57 | 9A57, 9A8C |
| 37 | 9A28 | 騨 | J0 | AJ | 9A52 | 9A52 |
| 38 | 9C2E | 鰮 | T3,J0,K1 | CJ | 9CC1 | 9C1B, 9CC1 |
| 39 | 9D0E | 鴎 | J0 | AJ | 9DD7 | 9DD7, 9E25 |
| 40 | 9D2C | 鴬 | J0 | AJ | 9DAF | 83BA, 9DAF |
| 41 | 9D8F | 鶏 | J0 | AJ | 96DE | 96DE, 9CEE, 9DC4, 9E21 |
| 42 | 9EBA | 麺 | J0 | AJ | 9EB5 | 9762, 9EAA, 9EB5 |

**Integration Panel recommendations:**

Following prior feedback from IP, these 42 characters are now defined as ‘out-of-repertoire’, which is a positive development. However, some work is still required in the variant set including these 42 characters/code points. This is explored in more detail in Sections 5.1.1.

# Variant considerations

## General

The IP mandate is to review each proposed LGR against the principles established in RFC 6912 as well as the process goals for the Root Zone LGR. As part of this, the IP aims to understand the foundation on which Variants have been assigned. Possible reasons include established practice in other zones of the DNS, documented linguistic rationale, or, for example, alignment with other Root Zone LGRs. This is the context in which the IP has analyzed CLGR10 against other sources. In this analysis, the dotAsia table is taken as a stand-in for existing practice, while the Unihan database represents publicly available data of a linguistic nature.

While neither is guaranteed to be “correct” and there are disagreements between these two, they can be broadly understood as reasonable starting points for determining what constitutes a variant. Where variant mappings in CLGR10 do not match either of these existing sources, the IP needs additional information to understand that the assigned variant mapping is in keeping with the principles, for example the Least Astonishment or Stability principles.

**Integration Panel recommendations:**

In general, the IP sees no problem with variant sets that are retained from previously deployed IDN tables. However, over the course of developing CLGR (up to its current draft version CLGR10), many changes were introduced that now create discrepancies to that earlier deployed practice.

A deviation from existing, deployed practice for Chinese variants may be motivated by a number of possible reasons. One of these reasons is clearly the need to share the Root Zone with a Korean LGR, and therefore pre-integration with KLGR can be an acceptable reason to split any variant set with at least two code points in the joint CLGR/KLGR Han subset.

However, for any of the other changes, there is the question of how confident the IP can be that these changes are stable, and not likely to receive pressure to overturn them at some point in the future. Section A.3.6 of the [Procedure] discusses this need for stability and the requirement to be highly confident in the details of a proposed RZ-LGR in the context of the "Principles to constrain the label generation rules".

The IP feels that for variant mappings that have changed, compared to deployed practice, it is essential to have a solid accounting of the motivation and cause for the change in order to be confident of the stability of the change. For variant rules that match long-standing practice, their longevity itself may serve to establish some confidence in their stability.

In order to evaluate the merits of these changes, the IP requires additional evidence and information on the rationale and justifications behind the CGP’s decisions. In some cases, the IP also suggests corrections or a different treatment altogether. The following summarizes the IP findings and requests:

1. The Integration have only found a limited set of issues with variant sets modified with the addition of 20 code points on top of dotAsia (the other 42 code points are out-of repertoire variants). Nevertheless, some of these require clarification.
2. Out-of-repertoire variants should have a reflexive mapping defined as ‘out-of-repertoire-var’ instead of ‘out-of-repertoire’, and **their mapping to other code points must be ‘blocked’.**
3. Variant sets that pertain to code points that are additions compared to the repertoires defined in the original Chinese repertoire (dotcn and dottw) and for which the variant relations are treated differently between dotAsia and CLGR10 **should be documented using referenced evidence**. **Alternatively, or in addition, any principles used in assigning these variants should be documented**.
4. The Integration Panel recognizes the necessary changes made to accommodate the pre-integration between the CGP and the KGP. While these changes introduce many differences in Chinese variant sets from previous practice (such as dotAsia, dotcn and dottw), the nature of the Root shared CJK variant sets makes this approach reasonable. However, **the IP found some anomalies in these changes that need to be either corrected or clarified.**
5. The Integration also found differences concerning variant sets between CLGR10 and the original Chinese repertoire found in dotcn and dottw. These differences should also be documented using referenced evidence. Alternatively, or in addition, **any principles used in assigning these variants should be documented.**

## Types of differences

In its analysis the IP notes that there are different types of discrepancies and differences between CLGR10 and the aforementioned sources. In some cases, variant sets in CLGR10 have additional members compared to already deployed IDN tables, such as dotAsia. In other cases, the variant sets are identical, but the mappings point to different code points. In the former case, the review needs to ensure that the mapping makes sense given the nature of additions. In the latter case, it should be possible to understand why CLGR10 would diverge from an existing IDN table which is already deployed and addresses roughly the same repertoire.

There is one natural source of difference between dotAsia (and the second-level IDN tables on which its development was based) and the Root Zone LGR. This difference is the need for the Root Zone to accommodate the use of Han Script for other communities, for example Korean. Where variant mappings result in in-repertoire variants for the KLGR, it is presumed that any changes from CLGR7 reflect the result of consensus between KPG and CGP, and that consensus represents justification for making those changes. Such justification, however, cannot be presumed in cases where only one code point in a variant relation is in the KLGR repertoire and no possible effects on in-repertoire KLGR variants are possible.

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

The following sections describe in detail the differences between CLGR10 and dotAsia. Where the variants set from the two LGRs differ in their definition but are overlapping, tables listing both variant sets are shown together and the header indications whether the set from CLGR10 or dotAsia comes first or second. Differences in variant mappings are highlighted in red. Although the comparison was mostly mechanically generated, some manual editing was necessary. Therefore, in case of discrepancies, the actual XML files are authoritative.

### Table format for variant sets

The format of each table follows the format used in the HTML-formatted LGR tables, 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 → direction and information for the reverse ← direction 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 “≡”.

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. Therefore, all variant sets are 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 CLGR10 and dotAsia are presented together, the CLGR10 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.

### Source information

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



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 reference 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 reference 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 J0 mapping would normally be considered “Japanese-specific”.

# Variant set differences by type of difference

## Additional repertoire (CLGR10 adds 62 code points not in dotAsia)

In most cases, the variant set in CLGR10 usually has a single additional code point compared to the same variant set defined in dotAsia, as well as the required additional mappings; all other mappings retain the same variant types. With additions affecting existing variant sets, changes are expected and at first reading most of these changes appear acceptable. 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.

These 62 new code positions came from 3 sources:

* 2 HKSCS characters that were left out when processing HKIRC request
* 18 characters from the Table of General Standard Chinese Characters (TGSCC formerly known as Normalized Hanzi List for Common Use (NHCU) in previous drafts of CLGR)
* 42 IICORE characters from J source and K sources part of Chinese based variant sets. These are also part of draft versions of the JGP and KGP repertoires and correspond to the out-of-repertoire variants. While CLGR10 has created a reflexive mapping of ‘out-of-repertoire’ (which should be changed to ‘out-of-repertoire-var’ for these code points, the XML file still contain mapping from and to these code points that are not of the value ‘blocked’. This needs to be corrected.

Of these 62 code points, one : U+48BC part of the 18 TGSCC set is a singleton, all others are part of variant sets. Finally, 7 of the following 62 entries combine context of additions to dotAsia (this section) with one of the following sections:

* **Code points added to CLGR10 through dotAsia but treated differently**, this concerns the entries for U+3A5C, U+58B5, U+7E4B, and U+9421.
* **Variants sets with pre-existing code points pre-integrated with KLGR**, this concerns the entry for U+9EB9.
* **Variant sets with pre-existing code points but different variant types**, this concerns the entry for U+61F4 and U+74A2.

These 7 entries are described in more details in section 5.1.4.

### Review of variant sets for out-of-repertoire variant beyond dotAsia – 42 sets

The following table shows these 42 code points that are specified in CLGR10 as out-of-repertoire-var that were added in previous versions of CLGR to complete the IICORE coverage but where originally from K or J sources and their corresponding variant sets. Except for the change to ‘blocked’ as mentioned above, most these variant sets are acceptable as presented. Some of these variant sets are ‘hybrid’ case in as such the differences result from multiple cause, these are further described in specific notes detailed in the section 5.1.4.

| **No** | **Code** | **CLGR10 Variant set glyphs and code points** | | | | | | | **Note** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 3960 | 㥠 3960 | 諝8ADD | 谞8C1E |  |  |  |  |  |
| 2 | 4FAD | 侭4FAD | 儘5118 | 尽5C3D | 盡76E1 |  |  |  |  |
| 3 | 51E6 | 処51E6 | 处5904 | 處8655 |  |  |  |  |  |
| 4 | 56A2 | 嚢56A2 | 囊56CA |  |  |  |  |  |  |
| 5 | 61F4 | 忏5FCF | 懴61F4 | 懺61FA |  |  |  |  | 1 |
| 6 | 6442 | 挕6315 | 摂6442 | 摄6444 | 攝651D |  |  |  |  |
| 7 | 663B | 昂6602 | 昻663B |  |  |  |  |  |  |
| 8 | 685C | 桜685C | 樱6A31 | 櫻6AFB |  |  |  |  |  |
| 9 | 685F | 栈6808 | 桟685F | 棧68E7 | 轏8F4F |  |  |  |  |
| 10 | 6D9C | 涜6D9C | 渎6E0E | 瀆7006 |  |  |  |  |  |
| 11 | 6E8C | 泼6CFC | 溌6E8C | 潑6F51 |  |  |  |  |  |
| 12 | 731F | 猎730E | 猟731F | 獵7375 |  |  |  |  |  |
| 13 | 784F | 揅63C5 | 研7814 | 硏784F |  |  |  |  |  |
| 14 | 7A36 | 稢7A22 | 稶7A36 |  |  |  |  |  |  |
| 15 | 7B86 | 箆7B86 | 篦7BE6 |  |  |  |  |  |  |
| 16 | 7C14 | 簑7C11 | 簔7C14 | 蓑84D1 |  |  |  |  |  |
| 17 | 7D9A | 続7D9A | 續7E8C | 续7EED |  |  |  |  |  |
| 18 | 7E4A | 孅5B45 | 縴7E34 | 繊7E4A | 纎7E8E | 纖7E96 | 纤7EA4 |  |  |
| 19 | 7E4B | 繋7E4B | 繫7E6B |  |  |  |  |  | 2 |
| 20 | 8133 | 脑8111 | 脳8133 | 腦8166 |  |  |  |  |  |
| 21 | 81D3 | 脏810F | 臓81D3 | 臟81DF | 髒9AD2 |  |  |  |  |
| 22 | 8217 | 舖8216 | 舗8217 | 鋪92EA | 铺94FA |  |  |  |  |
| 23 | 839F | 莟839F | 菡83E1 |  |  |  |  |  |  |
| 24 | 83B5 | 莵83B5 | 菟83DF |  |  |  |  |  |  |
| 25 | 86CD | 萤8424 | 蛍86CD | 螢87A2 |  |  |  |  |  |
| 26 | 8E99 | 躏8E8F | 躙8E99 | 躪8EAA |  |  |  |  |  |
| 27 | 9039 | 达8FBE | 迖8FD6 | 逹9039 | 達9054 |  |  |  |  |
| 28 | 91A4 | 酱9171 | 醤91A4 | 醬91AC |  |  |  |  |  |
| 29 | 91C8 | 釈91C8 | 释91CA | 釋91CB |  |  |  |  |  |
| 30 | 9421 | 鉄9244 | 銕9295 | 鐡9421 | 鐵9435 | 铁94C1 |  |  | 3 |
| 31 | 967A | 硷7877 | 碱78B1 | 礆7906 | 险9669 | 険967A | 險96AA | 鹼9E7C |  |
| 32 | 96B2 | 隲96B2 | 騭9A2D | 骘9A98 |  |  |  |  |  |
| 33 | 982C | 頬982C | 頰9830 | 颊988A |  |  |  |  |  |
| 34 | 98EE | 飮98EE | 飲98F2 | 饮996E |  |  |  |  |  |
| 35 | 9A12 | 騒9A12 | 騷9A37 | 骚9A9A |  |  |  |  |  |
| 36 | 9A13 | 験9A13 | 驗9A57 | 验9A8C |  |  |  |  |  |
| 37 | 9A28 | 騨9A28 | 驒9A52 |  |  |  |  |  |  |
| 38 | 9C2E | 鰛9C1B | 鰮9C2E | 鳁9CC1 |  |  |  |  |  |
| 39 | 9D0E | 鴎9D0E | 鷗9DD7 | 鸥9E25 |  |  |  |  |  |
| 40 | 9D2C | 莺83BA | 鴬9D2C | 鶯9DAF |  |  |  |  |  |
| 41 | 9D8F | 雞96DE | 鳮9CEE | 鶏9D8F | 鷄9DC4 | 鸡9E21 |  |  |  |
| 42 | 9EBA | 面9762 | 麪9EAA | 麵9EB5 | 麺9EBA |  |  |  |  |

### Review of variant sets for additional HKSCS characters beyond dotAsia – 2 sets

The following table shows variants sets modified by the addition of these 2 new HKSCS characters derived from HKIRC. These are called in specific notes detailed in the section 5.1.4.

| **No** | **Code** | **CLGR10 Variant set glyphs and code points** | | | | | | | | **Note** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 3A5C | 㩜 3A5C | 揽63FD | 攬652C |  |  |  |  |  | 4 |
| 2 | 58B5 | 坛575B | 埮57EE | 墰58B0 | 墵58B5 | 壇58C7 | 壜58DC | 罈7F48 | 罎7F4E | 5 |

### Review of variant sets for TGSCC characters beyond dotAsia – 18 sets

The following table corresponds to TGSCC addition (18 code points), 17 of them are in Appendix D 6.2.3.1-17 and 6.2.3.3.144. Missing in those appendix D items is U+9EB9 because of pre-integration coordination (appendix D 6.2.4-441). These are called in specific notes detailed in the section 5.1.4.

| **No** | **Code** | **Type** | **CLGR10 Variant set glyphs and code points** | | | | | | | **Note** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 48BC | N | 䢼 48BC |  |  |  |  |  |  |  |
| 2 | 5227 | V | 刦5226 | 刧5227 | 刼523C | 劫52AB |  |  |  |  |
| 3 | 524F | V | 创521B | 刱5231 | 剏524F | 剙5259 |  |  |  |  |
| 4 | 6060 | V | 怪602A | 恠6060 |  |  |  |  |  |  |
| 5 | 732F | N | 猯732F | 貒8C92 |  |  |  |  |  | 6 |
| 6 | 74A2 | V | 琉7409 | 瑠7460 | 璢74A2 |  |  |  |  | 7 |
| 7 | 750E | V | 甎750E | 砖7816 | 磚78DA |  |  |  |  |  |
| 8 | 754A | V | 畊754A | 耕8015 |  |  |  |  |  |  |
| 9 | 7ADA | V | 伫4F2B | 佇4F47 | 竚7ADA |  |  |  |  |  |
| 10 | 8262 | V | 樯6A2F | 檣6AA3 | 艢8262 |  |  |  |  |  |
| 11 | 88B5 | V | 衽887D | 袵88B5 |  |  |  |  |  |  |
| 12 | 894D | V | 杂6742 | 襍894D | 雑96D1 | 雜96DC | 雥96E5 |  |  |  |
| 13 | 8B0C | V | 歌6B4C | 謌8B0C |  |  |  |  |  |  |
| 14 | 8F19 | V | 輒8F12 | 輙8F19 | 辄8F84 |  |  |  |  |  |
| 15 | 945A | V | 鉆9246 | 鐉9409 | 鑚945A | 鑽947D | 钻94BB |  |  |  |
| 16 | 984B | V | 腮816E | 顋984B |  |  |  |  |  |  |
| 17 | 9DC0 | V | 鶿9DBF | 鷀9DC0 | 鹚9E5A |  |  |  |  |  |
| 18 | 9EB9 | N | 麴9EB4 | 麹9EB9 |  |  |  |  |  | 8 |

### Detailed review of variant sets modified by additions beyond dotAsia – 8 sets

Detailed review of variant sets modified by additions of characters beyond dotAsia is provided here. All but one corresponds to hybrid cases.

This variant set has one added member U+61F4 (from IICORE J source, out of repertoire variant). In addition, the mappings between U+61FA and U+5FCF have different type assignments between CLGR10 (1st) and dotAsia (2nd). As such, this is a hybrid case that combines this section with the section: ‘Variant sets with pre-existing code points but different variant types’.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 5FCF | 忏 | 5FCF | 忏 | ≡ | r-both |  | identity |
| 5FCF | 忏 | 61F4 | 懴 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 5FCF | 忏 | 61FA | 懺 | **→** | trad |  |  |
| **←** | simp |  |  |
| 61F4 | 懴 | 61F4 | 懴 | ≡ | out-of-repertoire |  | identity |
| 61F4 | 懴 | 61FA | 懺 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 61FA | 懺 | 61FA | 懺 | ≡ | r-trad |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 5FCF | 忏 | 5FCF | 忏 | ≡ | r-both |  | identity |
| 5FCF | 忏 | 61FA | 懺 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 61FA | 懺 | 61FA | 懺 | ≡ | r-trad |  | identity |

The mappings between U+61FA and U+5FCF in CLGR10 seem to follow the Unihan model (in it U+61FA is the Traditional Variant of 5FCF) and is therefore acceptable. However, all mapping to U+61F4 should be modified to ‘blocked’.

This variant set has one added member U+7E4B (not in dotAsia.) from the IICORE J source as out of repertoire variant. In addition, U+4FC2, U+7CFB and U+7E6B (all in dotAsia) are also included in the CLGR10 (1st and 2nd) and are mapped differently from dotAsia (3rd). This case is a hybrid of this category (one code point added not in dotAsia) and the next category (code points already in dotAsia but treated differently).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 4FC2 | 係 | 4FC2 | 係 | ≡ | r-trad |  | identity |
| 4FC2 | 係 | 7CFB | 系 | **→** | simp |  |  |
| **←** | trad |  |  |
| 7CFB | 系 | 7CFB | 系 | ≡ | r-both |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 7E4B | 繋 | 7E4B | 繋 | ≡ | out-of-repertoire |  | identity |
| 7E4B | 繋 | 7E6B | 繫 | ↔ | blocked |  |  |
| 7E6B | 繫 | 7E6B | 繫 | ≡ | r-both |  | identity |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 4FC2 | 係 | 4FC2 | 係 | ≡ | r-trad |  | identity |
| 4FC2 | 係 | 7CFB | 系 | **→** | simp |  |  |
| **←** | trad |  |  |
| 4FC2 | 係 | 7E6B | 繫 | ↔ | blocked |  |  |
| 7CFB | 系 | 7CFB | 系 | ≡ | r-both |  | identity |
| 7CFB | 系 | 7E6B | 繫 | **→** | trad |  |  |
| **←** | simp |  |  |
| 7E6B | 繫 | 7E6B | 繫 | ≡ | r-trad |  | identity |

In addition, U+4FC2, U+7CFB, and U+7E6B are part of KLGR (which explains the split between [4FC2, 7CFB] and 7E6B in the variant sets in CLGR10). As such, this change is acceptable.

This variant set has one added member U+9421 (not in dotAsia.) from IICORE, J source, out of repertoire. In addition, U+9244 (in dotAsia) from IICORE HKSCS was also included in the CLGR10 (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 | 銕 | ↔ | blocked |  |  |
| 9244 | 鉄 | 9421 | 鐡 | ↔ | blocked |  |  |
| 9244 | 鉄 | 9435 | 鐵 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 9244 | 鉄 | 94C1 | 铁 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 9295 | 銕 | 9295 | 銕 | ≡ | r-trad |  | identity |
| 9295 | 銕 | 9421 | 鐡 | ↔ | blocked |  |  |
| 9295 | 銕 | 9435 | 鐵 | ↔ | blocked |  |  |
| 9295 | 銕 | 94C1 | 铁 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 9421 | 鐡 | 9421 | 鐡 | ≡ | out-of-repertoire |  | identity |
| 9421 | 鐡 | 9435 | 鐵 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 9421 | 鐡 | 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 | 鉄 | 9244 | 鉄 | ≡ | r-trad |  | identity |
| 9244 | 鉄 | 9295 | 銕 | ↔ | blocked |  |  |
| 9244 | 鉄 | 9435 | 鐵 | ↔ | blocked |  |  |
| 9244 | 鉄 | 94C1 | 铁 | **→** | simp |  |  |
| **←** | trad |  |  |
| 9295 | 銕 | 9295 | 銕 | ≡ | r-trad |  | identity |
| 9295 | 銕 | 9435 | 鐵 | ↔ | blocked |  |  |
| 9295 | 銕 | 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 mapping for U+9421 to other code points should be all ‘blocked’. The differences concerning U+9244 mappings are not addressed at this point (Unihan has Semantic Variant mappings between U+9244, U+9295, and U+9435).

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 CLGR10 variant sets (1st and 2nd tables) and are mapped differently from dotAsia (3rd table). This case is a hybrid of this category (one code point added not in dotAsia) and of the next categories (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 | 㧛 | 39DB | 㧛 | ≡ | r-both |  | identity |
| 39DB | 㧛 | 64E5 | 擥 | ↔ | blocked |  |  |
| 64E5 | 擥 | 64E5 | 擥 | ↔ | r-both |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 3A5C | 㩜 | 3A5C | 㩜 | ≡ | r-both |  | identity |
| 3A5C | 㩜 | 63FD | 揽 | ↔ | blocked |  |  |
| 3A5C | 㩜 | 652C | 攬 | ↔ | blocked |  |  |
| 63FD | 揽 | 63FD | 揽 | ≡ | r-simp |  | identity |
| 63FD | 揽 | 652C | 攬 | **→** | trad |  |  |
| **←** | simp |  |  |
| 652C | 攬 | 652C | 攬 | ≡ | r-trad |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 39DB | 㧛 | 39DB | 㧛 | ≡ | r-simp |  | identity |
| 39DB | 㧛 | 63FD | 揽 | ↔ | blocked |  |  |
| 39DB | 㧛 | 64E5 | 擥 | **→** | trad |  |  |
| **←** | simp |  |  |
| 39DB | 㧛 | 652C | 攬 | ↔ | blocked |  |  |
| 63FD | 揽 | 63FD | 揽 | ≡ | r-simp |  | identity |
| 63FD | 揽 | 64E5 | 擥 | ↔ | blocked |  |  |
| 63FD | 揽 | 652C | 攬 | **→** | trad |  |  |
| **←** | simp |  |  |
| 64E5 | 擥 | 64E5 | 擥 | ≡ | r-trad |  | identity |
| 64E5 | 擥 | 652C | 攬 | ↔ | blocked |  |  |
| 652C | 攬 | 652C | 攬 | ≡ | r-trad |  | identity |

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



Unihan indicates that it is a semantic variant of U+652C 攬and U+64E5 擥. In addition U+3A2B 㨫(not in CLGR10) 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).

This variant set has one added member U+58B5 (not in dotAsia.) from the HKSCS set coming from HKIRC. In addition, U+58DC (in dotAsia) is also included in the CLGR10 (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 CLGR10 table add them as ‘blocked’ variants of all other members.

The CLGR10 set is an 8 members variant set.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 575B | 坛 | 575B | 坛 | ≡ | r-simp |  | identity |
| 575B | 坛 | 57EE | 埮 | ↔ | blocked |  |  |
| 575B | 坛 | 58B0 | 墰 | ↔ | blocked |  |  |
| 575B | 坛 | 58B5 | 墵 | ↔ | blocked |  |  |
| 575B | 坛 | 58C7 | 壇 | **→** | trad |  |  |
| **←** | simp |  |  |
| 575B | 坛 | 58DC | 壜 | ↔ | blocked |  |  |
| 575B | 坛 | 7F48 | 罈 | **→** | trad |  |  |
| **←** | simp |  |  |
| 575B | 坛 | 7F4E | 罎 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 57EE | 埮 | 57EE | 埮 | ≡ | r-both |  | identity |
| 57EE | 埮 | 58B0 | 墰 | ↔ | blocked |  |  |
| 57EE | 埮 | 58B5 | 墵 | ↔ | blocked |  |  |
| 57EE | 埮 | 58C7 | 壇 | ↔ | blocked |  |  |
| 57EE | 埮 | 58DC | 壜 | ↔ | blocked |  |  |
| 57EE | 埮 | 7F48 | 罈 | ↔ | blocked |  |  |
| 57EE | 埮 | 7F4E | 罎 | ↔ | blocked |  |  |
| 58B0 | 墰 | 58B0 | 墰 | ≡ | r-both |  | identity |
| 58B0 | 墰 | 58B5 | 墵 | ↔ | blocked |  |  |
| 58B0 | 墰 | 58C7 | 壇 | ↔ | blocked |  |  |
| 58B0 | 墰 | 58DC | 壜 | ↔ | blocked |  |  |
| 58B0 | 墰 | 7F48 | 罈 | ↔ | blocked |  |  |
| 58B0 | 墰 | 7F4E | 罎 | ↔ | blocked |  |  |
| 58B5 | 墵 | 58B5 | 墵 | ≡ | r-both |  | identity |
| 58B5 | 墵 | 58C7 | 壇 | ↔ | blocked |  |  |
| 58B5 | 墵 | 58DC | 壜 | ↔ | blocked |  |  |
| 58B5 | 墵 | 7F48 | 罈 | ↔ | blocked |  |  |
| 58B5 | 墵 | 7F4E | 罎 | ↔ | blocked |  |  |
| 58C7 | 壇 | 58C7 | 壇 | ≡ | r-trad |  | identity |
| 58C7 | 壇 | 58DC | 壜 | ↔ | blocked |  |  |
| 58C7 | 壇 | 7F48 | 罈 | ↔ | blocked |  |  |
| 58C7 | 壇 | 7F4E | 罎 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 58DC | 壜 | 58DC | 壜 | ≡ | r-both |  | identity |
| 58DC | 壜 | 7F48 | 罈 | ↔ | blocked |  |  |
| 58DC | 壜 | 7F4E | 罎 | ↔ | blocked |  |  |
| 7F48 | 罈 | 7F48 | 罈 | ≡ | r-trad |  | identity |
| 7F48 | 罈 | 7F4E | 罎 | ↔ | blocked |  |  |
| 7F4E | 罎 | 7F4E | 罎 | ≡ | 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 | 壜 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 575B | 坛 | 7F48 | 罈 | **→** | trad |  |  |
| **←** | simp |  |  |
| 575B | 坛 | 7F4E | 罎 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 58C7 | 壇 | 58C7 | 壇 | ≡ | r-trad |  | identity |
| 58C7 | 壇 | 58DC | 壜 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 58C7 | 壇 | 7F48 | 罈 | ↔ | blocked |  |  |
| 58C7 | 壇 | 7F4E | 罎 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 58DC | 壜 | 58DC | 壜 | ≡ | r-both |  | identity |
| 58DC | 壜 | 7F48 | 罈 | ↔ | blocked |  |  |
| 58DC | 壜 | 7F4E | 罎 | ↔ | blocked |  |  |
| 7F48 | 罈 | 7F48 | 罈 | ≡ | r-trad |  | identity |
| 7F48 | 罈 | 7F4E | 罎 | ↔ | blocked |  |  |
| 7F4E | 罎 | 7F4E | 罎 | ≡ | r-neither |  | identity |

The new CLGR10 corresponds to the recommendation that was made by IP on earlier feedback and as such, the result may be satisfactory. However, the fact that this creates another 8-member variant set with multiple allocatable mappings is not a situation that is reassuring. In addition, Unihan, while having a variant set including 6 of the members of this variant, does not include the new member U+58B5, nor U+7F4E. This variant set may need to be revisited.

This new variant set is created by combining the new U+732F from TGSCC 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 | 猯 | 732F | 猯 | ≡ | r-both |  | identity, reflexive |
| 732F | 猯 | 8C92 | 貒 | ↔ | blocked |  |  |
| 8C92 | 貒 | 8C92 | 貒 | ≡ | r-both |  | identity, reflexive |

The code point U+732F has G, T, J, and K sources and is part of the TGSCC.



Unihan does not define any variant for this character; therefore no conclusion can be made.

This variant set has one added member U+74A2 from TGSCC. In addition, the mapping between U+7409 and U+7460 is different between CLGR10 (1st) and dotAsia (2nd). As such, this is an hybrid case combining this section with the section: ‘Variant sets with pre-existing code points but different variant types’.

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



Unihan kSemanticVariant field indicates that this is a variant of U+7409 琉. More information is needed to evaluate the difference of mapping for U+7409 and U+7460.

The character U+9EB9 from TGSCC (also IICORE J source) is associated with the code point U+9EB4 which is part of the variant set U+66F2, U+9EAF, and U+9EB4 in dotAsia (3rd table). In CLGR10, U+66F2 and U+9EAF are in a variant set (1st table), and the pair U+9EB4, U+9EB9 is in another variant set (2nd table). This is a hybrid case with the KLGR pre-integration.

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



Unihan has variant relationship for U+9EAF, U+9EB4, and U+9EB9, but not U+66F2. Code points U+66F2 and U+9EB4 are part of KLGR.

## Code points added to CLGR10 through dotAsia but treated differently

The variant sets in this section have code points that originate from the dotAsia repertoire, but the chosen variant mappings differ from those used in the original dotAsia set. To show the differences, variant sets may be listed twice, once for CLGR10 and once for dotAsia.

At this point the IP lacks the information to either accept or reject the mappings found in CLGR10 and detailed here. While CLGR10 provides additional information and documentation concerning these differences (see section 6.2.2 and 6.2.3 of [Proposal], neither the Appendix D or Appendix E provide usable reference to determine the mapping used for the variant sets shown as problematic is this section (5.2).

For example, code point U+3A18 is mentioned in one line in Appendix D 6.2.2-126 as another variant of U+64E4 (but does not have its own line); it has its own line in Appendix D 6.2.3-144, but in that table U+64E4 does not have its own line. To reconstruct the reference, one has to merge both tables. In addition, while the Integration Panel expected to find the differences between dotAsia and CLGR10 documented in these 2 tables (Appendix D 6.2.2-126 and Appendix D 6.2.3-144) it could not detect all of them.

In addition to the 45 entries part of this section, there are 4 entries from the previous section corresponding to the addition of code points U+3A5C, U+58B5, U+7E4B, and U+9421. These need the same documentation because they also contain differences related to this section.

**Integration Panel recommendation:**

Please provide the information concerning these changes in a format that can be used by the IP, all tables should be transitive (one line per variant set element). In addition, when the changes introduced by CLGR10 contradicts Unihan or are not supported by Unihan (about 31+4 variant sets), additional rationale should be provided.

### Summary of the list of additions treated differently – 45 items

Note that one entry including the code point: U+637F is affected by the KLGR pre-integration.

Legend:

* Code point in red represent dotAsia additions,
* Green-highlighted code points represents either singleton or additional variant sets,
* Digits (1 to 31) in the last column indicates indexes in the following sub-section in 5.2.2 and corresponds to cases that need further investigation. If the last column is blank, it means that the CLGR10 modification is acceptable.

| **No** | **Variant set(s) glyphs and code points** | | | | | | | | | **Note** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 㨘3A18 | 擤64E4 |  |  |  |  |  |  |  | 1 |
| 2 | 㩒3A52 | 擒64D2 |  |  |  |  |  |  |  | 2 |
| 3 | 䱽4C7D | 䲝4C9D | 鯧9BE7 | 鲳9CB3 |  |  |  |  |  |  |
| 4 | 假5047 | 叚53DA | 段6BB5 |  |  |  |  |  |  | 3 |
| 5 | 净51C0 | 凈51C8 | 浄6D44 | 淨6DE8 | 瀞701E |  |  |  |  |  |
| 6 | 刊520A | 刋520B | 栞681E |  |  |  |  |  |  |  |
| 7 | 劤52A4 | 劲52B2 | 勁52C1 |  |  |  |  |  |  |  |
| 8 | 发53D1 | 彂5F42 | 発767A | 發767C | 髪9AEA | 髮9AEE |  |  |  |  |
| 9 | 吻543B | 呅5445 | 呡5461 | 脗8117 |  |  |  |  |  | 4 |
| 10 | 园56ED | 園5712 | 薗8597 |  |  |  |  |  |  | 5 |
| 11 | 宝5B9D | 寳5BF3 | 寶5BF6 | 珤73E4 |  |  |  |  |  | 6 |
| 12 | 帚5E1A | 箒7B92 | 菷83F7 |  |  |  |  |  |  | 7 |
| 13 | 弄5F04 | 挵6335 | 衖8856 |  |  |  |  |  |  | 8 |
| 14 | 总603B | 惣60E3 | 揔63D4 | 捴6374 | 搃6403 | 摠6460 | 総7DCF | 縂7E02 | 總7E3D | 9 |
| 15 | 慼617C | 慽617D | 戚621A |  |  |  |  |  |  |  |
| 16 | 扬626C | 揚63DA | 敭656D | 颺98BA | 飏98CF |  |  |  |  | 10 |
| 17 | 抬62AC | 擡64E1 |  |  |  |  |  |  |  | 11 |
| 18 | 捿637F | 栖6816 | 棲68F2 |  |  |  |  |  |  | 12 |
| 19 | 昞661E | 昺663A | 炳70B3 |  |  |  |  |  |  |  |
| 20 | 晧6667 | 暠66A0 | 皓7693 | 皜769C |  |  |  |  |  |  |
| 21 | 椀6900 | 盌76CC | 碗7897 | 鋺92FA |  |  |  |  |  | 13 |
| 22 | 泷6CF7 | 滝6EDD | 瀧7027 |  |  |  |  |  |  | 14 |
| 23 | 灶7076 | 竃7AC3 | 竈7AC8 |  |  |  |  |  |  | 15 |
| 24 | 炼70BC | 煉7149 | 錬932C | 鍊934A |  |  |  |  |  |  |
| 25 | 爀7200 | 赫8D6B |  |  |  |  |  |  |  | 16 |
| 26 | 獏734F | 貘8C98 |  |  |  |  |  |  |  | 17 |
| 27 | 珡73E1 | 琴7434 |  |  |  |  |  |  |  |  |
| 28 | 琍740D | 璃7483 | 瓈74C8 |  |  |  |  |  |  | 18 |
| 29 | 畺757A | 疆7586 |  |  |  |  |  |  |  | 19 |
| 30 | 痫75EB | 癇7647 | 癎764E |  |  |  |  |  |  | 20 |
| 31 | 睹7779 | 覩89A9 |  |  |  |  |  |  |  | 21 |
| 32 | 秕79D5 | 粃7C83 |  |  |  |  |  |  |  | 22 |
| 33 | 筯7B6F | 箸7BB8 |  |  |  |  |  |  |  | 23 |
| 34 | 綿7DBF | 緜7DDC | 绵7EF5 |  |  |  |  |  |  | 24 |
| 35 | 肕8095 | 靭976D | 靱9771 | 韌97CC | 韧97E7 |  |  |  |  | 25 |
| 36 | 膳81B3 | 饍994D |  |  |  |  |  |  |  | 26 |
| 37 | 萌840C | 萠8420 | 蕄8544 |  |  |  |  |  |  | 27 |
| 38 | 蓚84DA | 蓨84E8 |  |  |  |  |  |  |  | 28 |
| 39 | 転8EE2 | 轉8F49 | 转8F6C |  |  |  |  |  |  |  |
| 40 | 边8FB9 | 辺8FBA | 邉9089 | 邊908A |  |  |  |  |  |  |
| 41 | 鉄9244 | 銕9295 | 鐡9421 | 鐵9435 | 铁94C1 |  |  |  |  | 29 |
| 42 | 飆98C6 | 飇98C7 | 飈98C8 | 飙98D9 | 飚98DA |  |  |  |  | 30 |
| 43 | 駅99C5 | 驛9A5B | 驿9A7F |  |  |  |  |  |  |  |
| 44 | 鴟9D1F | 鵄9D44 | 鸱9E31 |  |  |  |  |  |  | 31 |
| 45 | 齡9F61 | 齢9F62 | 龄9F84 |  |  |  |  |  |  |  |

### Detailed list of differences for dotAsia additions – 31 sets

The following numbered items list all the information available to the IP at this point for each of the differences found. In some cases, that information appears to contradict a choice made in CLGR10. For these cases it is especially important to reference the information on the basis of which the CGP made its determination. In other cases, the available information is inconclusive as far as the IP can determine. Where the data available to the IP is inconclusive, additional clarification from the CGP is needed.

In order to evaluate the merits of these different choices the IP requests the GCP to provide documentation of the rationale, background, references etc. that motivate EACH of the chosen mappings. Note that this request applies to all items, even though it is not stated for each.

The code point U+3A18 was included in CLGR10 because of its membership in IICORE HKSCS but is has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

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

Unihan in the entry for U+3A18 mentions in its kDefinition field that it is “(same as U+64E4 擤) to blow the nose with the fingers; (Cant.) to scour; to rebuke; to hit with a ball”; but it has no traditional/simplified mapping. Therefore, while a semantic variant is implied, the traditional/simplified mapping added by CLGR10 is not supported.

The code point U+3A52 was included in CLGR10 because of its membership in IICORE HKSCS but is treated differently between CLGR10 (table follows) and dotAsia (where it is a singleton reflexive variant of ‘r-both’).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 3A52 | 㩒 | 3A52 | 㩒 | ≡ | r-trad |  | identity |
| 3A52 | 㩒 | 64D2 | 擒 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 64D2 | 擒 | 64D2 | 擒 | ≡ | r-both |  | identity |

The simplified mapping between U+3A52 and U+64D2 in CLGR10 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 捦, not supported by either CLGR10 or dotAsia.

The code point U+53DA was included in CLGR10 because of its membership in IICORE HKSCS but is treated differently between CLGR10 (1st) and dotAsia (2nd). In one case, it is a variant of U+6BB5, in the other a variant of U+5047. Code points U+6BB5 and U+5047 are members of both CLGR10 and dotAsia with ‘r-both’ mapping.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 53DA | 叚 | 53DA | 叚 | ≡ | r-both |  | identity |
| 53DA | 叚 | 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 | 叚 | ↔ | blocked |  |  |
| 53DA | 叚 | 53DA | 叚 | ≡ | r-both |  | identity |

Unihan does not bring any clarification either way (it describes a variant relationship between U+5047 and U+4EEE part of CLGR10 but not included here). More information is required to evaluate this. Code points U+5047 and U+6BB5 are part of KLGR.

The code point U+8117 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 543B | 吻 | 543B | 吻 | ≡ | r-both |  | identity |
| 543B | 吻 | 5445 | 呅 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 543B | 吻 | 5461 | 呡 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 543B | 吻 | 8117 | 脗 | **→** | blocked |  |  |
| **←** | both |  |  |
| 5445 | 呅 | 5445 | 呅 | ≡ | r-trad |  | identity |
| 5445 | 呅 | 5461 | 呡 | ↔ | blocked |  |  |
| 5445 | 呅 | 8117 | 脗 | ↔ | blocked |  |  |
| 5461 | 呡 | 5461 | 呡 | ≡ | r-trad |  | identity |
| 5461 | 呡 | 8117 | 脗 | ↔ | blocked |  |  |
| 8117 | 脗 | 8117 | 脗 | ≡ | r-neither |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 543B | 吻 | 543B | 吻 | ≡ | r-both |  | identity |
| 543B | 吻 | 5445 | 呅 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 543B | 吻 | 5461 | 呡 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 543B | 吻 | 8117 | 脗 | ↔ | blocked |  |  |
| 5445 | 呅 | 5445 | 呅 | ≡ | r-trad |  | identity |
| 5445 | 呅 | 5461 | 呡 | ↔ | blocked |  |  |
| 5445 | 呅 | 8117 | 脗 | ↔ | blocked |  |  |
| 5461 | 呡 | 5461 | 呡 | ≡ | r-trad |  | identity |
| 5461 | 呡 | 8117 | 脗 | ↔ | blocked |  |  |
| 8117 | 脗 | 8117 | 脗 | ≡ | r-both |  | identity |

Unihan has a mapping between U+543B and U+5461 (and U+80B3 not part of this variant set), but not U+5445 or U+8117. More information is needed to evaluate this.

The code point U+8597 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (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 | 薗 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 5712 | 園 | 5712 | 園 | ≡ | r-trad |  | identity |
| 5712 | 園 | 8597 | 薗 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 8597 | 薗 | 8597 | 薗 | ≡ | 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 | 薗 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 5712 | 園 | 5712 | 園 | ≡ | r-trad |  | identity |
| 5712 | 園 | 8597 | 薗 | ↔ | blocked |  |  |
| 8597 | 薗 | 8597 | 薗 | ≡ | r-trad |  | identity |

Unihan has a variant relationship between U+56ED and U+5712 (and a kSemanticVariant mapping to U+5313 not in this variant set), but no mapping to U+8597. More information is needed to evaluate this.

The code point U+73E4 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (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 | 寳 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 5B9D | 宝 | 5BF6 | 寶 | **→** | trad |  |  |
| **←** | simp |  |  |
| 5B9D | 宝 | 73E4 | 珤 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 5BF3 | 寳 | 5BF3 | 寳 | ≡ | r-neither |  | identity |
| 5BF3 | 寳 | 5BF6 | 寶 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 5BF3 | 寳 | 73E4 | 珤 | ↔ | blocked |  |  |
| 5BF6 | 寶 | 5BF6 | 寶 | ≡ | r-trad |  | identity |
| 5BF6 | 寶 | 73E4 | 珤 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 73E4 | 珤 | 73E4 | 珤 | ≡ | r-neither |  | identity |

Unihan has a variant set for U+5B9D, U+5BF3, and U+5BF6, not including U+73E4. This supports the dotAsia mapping, not the CLGR10 set.

The code point U+7B92 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 5E1A | 帚 | 5E1A | 帚 | ≡ | r-both |  | identity |
| 5E1A | 帚 | 7B92 | 箒 | **→** | blocked |  |  |
| **←** | both |  |  |
| 5E1A | 帚 | 83F7 | 菷 | **→** | blocked |  |  |
| **←** | both |  |  |
| 7B92 | 箒 | 7B92 | 箒 | ≡ | r-neither |  | identity |
| 7B92 | 箒 | 83F7 | 菷 | ↔ | blocked |  |  |
| 83F7 | 菷 | 83F7 | 菷 | ≡ | r-neither |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 5E1A | 帚 | 5E1A | 帚 | ≡ | r-both |  | identity |
| 5E1A | 帚 | 7B92 | 箒 | ↔ | blocked |  |  |
| 5E1A | 帚 | 83F7 | 菷 | **→** | blocked |  |  |
| **←** | both |  |  |
| 7B92 | 箒 | 7B92 | 箒 | ≡ | r-both |  | identity |
| 7B92 | 箒 | 83F7 | 菷 | ↔ | blocked |  |  |
| 83F7 | 菷 | 83F7 | 菷 | ≡ | r-neither |  | identity |

Unihan has a variant for all these code points. It indicates that U+5E1A is a semantic variant of U+7B92. Therefore, the correlation is established, but not the type of preferred mapping or the reason why there is a difference between CLGR10 and dotAsia.

The code point U+6335 was included in CLGR10 because of its membership in IICORE HKSCS set but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

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

Unihan does not bring any clarification (U+5F04 has variant relationship with U+7627 and U+759F which is a separate CLGR10 variant set, but U+6335 and U+8856 have no variant in Unihan); More information is needed to evaluate this.

The code point U+60E3 was included in CLGR10 because of its membership in IICORE HKSCS; however it is treated differently. In dotAsia, it is part of another variant set with U+63D4. This case is a mixed of added code point (U+60E3) and re-arrangement of variant sets due to pre-integration with the KLGR which is reviewed in another section.

In dotAsia, U+603B, U+6374, U+6403, U+6460, U+7DCF, U+7E02, and U+7E3D have a variant relationship (3rd table), and the pair U+60E3-U+63D4 have another variant relationship (4th table).

In CLGR10, U+603B, U+7DCF, U+7E02 and U+7E3D are a variant set (1st table), the U+60E3, U+6374, U+6403, and U+66460 are part of another variant set (2nd table) and the remaining code point: U+63D4 is a singleton reflexive variant of type ‘r-both’.

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

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 603B | 总 | 603B | 总 | ≡ | r-simp |  | identity |
| 603B | 总 | 6374 | 捴 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 603B | 总 | 6403 | 搃 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 603B | 总 | 6460 | 摠 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 603B | 总 | 7DCF | 総 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 603B | 总 | 7E02 | 縂 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 603B | 总 | 7E3D | 總 | **→** | trad |  |  |
| **←** | simp |  |  |
| 6374 | 捴 | 6374 | 捴 | ≡ | r-neither |  | identity |
| 6374 | 捴 | 6403 | 搃 | ↔ | blocked |  |  |
| 6374 | 捴 | 6460 | 摠 | ↔ | blocked |  |  |
| 6374 | 捴 | 7DCF | 総 | ↔ | blocked |  |  |
| 6374 | 捴 | 7E02 | 縂 | ↔ | blocked |  |  |
| 6374 | 捴 | 7E3D | 總 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 6403 | 搃 | 6403 | 搃 | ≡ | r-neither |  | identity |
| 6403 | 搃 | 6460 | 摠 | ↔ | blocked |  |  |
| 6403 | 搃 | 7DCF | 総 | ↔ | blocked |  |  |
| 6403 | 搃 | 7E02 | 縂 | ↔ | blocked |  |  |
| 6403 | 搃 | 7E3D | 總 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 6460 | 摠 | 6460 | 摠 | ≡ | r-trad |  | identity |
| 6460 | 摠 | 7DCF | 総 | ↔ | blocked |  |  |
| 6460 | 摠 | 7E02 | 縂 | ↔ | blocked |  |  |
| 6460 | 摠 | 7E3D | 總 | ↔ | blocked |  |  |
| 7DCF | 総 | 7DCF | 総 | ≡ | r-neither |  | identity |
| 7DCF | 総 | 7E02 | 縂 | ↔ | blocked |  |  |
| 7DCF | 総 | 7E3D | 總 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 7E02 | 縂 | 7E02 | 縂 | ≡ | r-neither |  | identity |
| 7E02 | 縂 | 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 | 揔 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 63D4 | 揔 | 63D4 | 揔 | ≡ | r-both |  | identity |

Unihan has a variant relationship between U+603B, U+6374, U+6403, U+6460, U+7DCF, U+7E02, and U+7E3D (U+603B and U+7E3D are trad-simp), but has no variant mapping for U+60E3 or U+63D4; More information is needed to evaluate this.

However, code points U+6460 and U+7E3D are part of KLGR, which explains the split of the original 7-member variant set.

The code point U+656D was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd). Note that U+656D is mostly used in a Korean context.

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

Unihan has simp-trad variant relationship between U+626C and U+63DA, and another relationship between U+98BA and U+98CF, and none for U+656D. No conclusion can be made unless additional documentation is provided.

The code point U+64E1 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (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 unless more documentation is provided.

The code point U+637F was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between dotAsia and CLGR10. In CLGR10 (1st table) it is part of a variant relationship with U+68F2, the other code point is U+6816 is a singleton reflexive variant of type ‘r-both’. In dotAsia, U+637F is a singleton reflexive variant of type ‘r-both’ and the two other code points: U+6816 and U+68F2 (2nd table) are in a variant relationship.

This is a hybrid case, because it is also affected by KLGR pre-integration (separation of U+6816 and U+68F2).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 637F | 捿 | 637F | 捿 | ≡ | r-simp |  | identity |
| 637F | 捿 | 68F2 | 棲 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 68F2 | 棲 | 68F2 | 棲 | ≡ | r-both |  | identity |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 6816 | 栖 | 6816 | 栖 | ≡ | r-both |  | identity |
| 6816 | 栖 | 68F2 | 棲 | **→** | trad |  |  |
| **←** | simp |  |  |
| 68F2 | 棲 | 68F2 | 棲 | ≡ | r-trad |  | identity |

Unihan has a trad-simp variant relationship between U+6816 and U+68F2, but nothing for 637F. This would support the dotAsia mappings. All code points are part of KLGR (which has a blocked mapping between U+637F and U+68F2). The KLGR integration split the pair U+6816-U+68F2. Unclear what the best solution should be here (may be have 3 singletons?).

The code points U+6900 and U+76CC were included in CLGR10 because of their IICORE HKSCS property but are treated differently between CLGR10 (1st) and dotAsia (2nd).

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

While in Unihan there is a kSemanticVariant mapping between U+6900 and U+7897, the ‘trad’ mapping suggested by CLGR10 seems dubious, and the dotAsia mapping seems preferable. The other differences need more information to evaluate them.

The code point U+6EDD was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (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 CLGR10. More information is needed to evaluate this.

The code point U+7AC3 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

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

Unihan has a variant relationship between all code points; more documentation is needed to evaluate this.

The code point U+7200 was included in CLGR10 because of its membership in IICORE HKSCS but is treated differently between dotAsia and CLGR10 (table follows). In dotAsia both code points are singleton reflexive variants of type ‘r-both’.

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

Unihan does not have a variant relationship between these code points. However they have the same kCantonese data: ‘haak1’ ; More information is needed to evaluate this. These 2 code points are part of KLGR and in a ‘blocked’ variant relationship in KLGR.

The code point U+734F was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

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

Unihan has Zvariant relationship between these code points and U+8C8A; More information is needed to evaluate this.

The code point U+74C8 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (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 | 瓈 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 7483 | 璃 | 7483 | 璃 | ≡ | r-both |  | identity |
| 7483 | 璃 | 74C8 | 瓈 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 74C8 | 瓈 | 74C8 | 瓈 | ≡ | 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 | 瓈 | ↔ | blocked |  |  |
| 7483 | 璃 | 7483 | 璃 | ≡ | r-both |  | identity |
| 7483 | 璃 | 74C8 | 瓈 | ↔ | blocked |  |  |
| 74C8 | 瓈 | 74C8 | 瓈 | ≡ | r-both |  | identity |

Although Unihan describes a kSemanticVariant between U+74C8 and U+7483, and U+740D shares the kCantonese value of ‘lei4’ with the two other code points. However, there is no traditional or simplified mapping implied for those code points. More information is needed to evaluate this.

The code point U+757A was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 757A | 畺 | 757A | 畺 | ≡ | r-neither |  | identity |
| 757A | 畺 | 7586 | 疆 | **→** | both |  |  |
| **←** | blocked |  |  |
| 7586 | 疆 | 7586 | 疆 | ≡ | r-both |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 757A | 畺 | 757A | 畺 | ≡ | r-both |  | identity |
| 757A | 畺 | 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’. This change needs to be justified.

The code point U+764E was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

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

Unihan has a trad-simp mapping between U+75EB and U+7647 and they all share a kCantonese value of ‘haan4’; More information is needed to evaluate this.

The code point U+89A9 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (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’. This change needs to be justified.

The code point U+7C83 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 79D5 | 秕 | 79D5 | 秕 | ≡ | r-both |  | identity |
| 79D5 | 秕 | 7C83 | 粃 | **→** | blocked |  |  |
| **←** | both |  |  |
| 7C83 | 粃 | 7C83 | 粃 | ≡ | r-neither |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 79D5 | 秕 | 79D5 | 秕 | ≡ | r-both |  | identity |
| 79D5 | 秕 | 7C83 | 粃 | ↔ | blocked |  |  |
| 7C83 | 粃 | 7C83 | 粃 | ≡ | 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’. This change needs to be justified.

The code point U+7B6F was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 7B6F | 筯 | 7B6F | 筯 | ≡ | r-neither |  | identity |
| 7B6F | 筯 | 7BB8 | 箸 | **→** | both |  |  |
| **←** | blocked |  |  |
| 7BB8 | 箸 | 7BB8 | 箸 | ≡ | r-both |  | identity, reflexive |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 7B6F | 筯 | 7B6F | 筯 | ≡ | r-both |  | identity |
| 7B6F | 筯 | 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’. This change needs to be justified.

The code point U+7DDC was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 7DBF | 綿 | 7DBF | 綿 | ≡ | r-trad |  | identity |
| 7DBF | 綿 | 7DDC | 緜 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 7DBF | 綿 | 7EF5 | 绵 | **→** | simp |  |  |
| **←** | trad |  |  |
| 7DDC | 緜 | 7DDC | 緜 | ≡ | r-neither |  | identity |
| 7DDC | 緜 | 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 | 緜 | ↔ | blocked |  |  |
| 7DBF | 綿 | 7EF5 | 绵 | **→** | simp |  |  |
| **←** | trad |  |  |
| 7DDC | 緜 | 7DDC | 緜 | ≡ | r-both |  | identity |
| 7DDC | 緜 | 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’. This change needs to be justified.

The code point U+9771 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 8095 | 肕 | 8095 | 肕 | ≡ | r-trad |  | identity |
| 8095 | 肕 | 976D | 靭 | ↔ | blocked |  |  |
| 8095 | 肕 | 9771 | 靱 | ↔ | blocked |  |  |
| 8095 | 肕 | 97CC | 韌 | ↔ | blocked |  |  |
| 8095 | 肕 | 97E7 | 韧 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 976D | 靭 | 976D | 靭 | ≡ | r-neither |  | identity |
| 976D | 靭 | 9771 | 靱 | ↔ | blocked |  |  |
| 976D | 靭 | 97CC | 韌 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 976D | 靭 | 97E7 | 韧 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 9771 | 靱 | 9771 | 靱 | ≡ | r-neither |  | identity |
| 9771 | 靱 | 97CC | 韌 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 9771 | 靱 | 97E7 | 韧 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 97CC | 韌 | 97CC | 韌 | ≡ | r-trad |  | identity |
| 97CC | 韌 | 97E7 | 韧 | **→** | simp |  |  |
| **←** | trad |  |  |
| 97E7 | 韧 | 97E7 | 韧 | ≡ | r-simp |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 8095 | 肕 | 8095 | 肕 | ≡ | r-trad |  | identity |
| 8095 | 肕 | 976D | 靭 | ↔ | blocked |  |  |
| 8095 | 肕 | 9771 | 靱 | ↔ | blocked |  |  |
| 8095 | 肕 | 97CC | 韌 | ↔ | blocked |  |  |
| 8095 | 肕 | 97E7 | 韧 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 976D | 靭 | 976D | 靭 | ≡ | r-neither |  | identity |
| 976D | 靭 | 9771 | 靱 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 976D | 靭 | 97CC | 韌 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 976D | 靭 | 97E7 | 韧 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 9771 | 靱 | 9771 | 靱 | ≡ | r-neither |  | identity |
| 9771 | 靱 | 97CC | 韌 | ↔ | blocked |  |  |
| 9771 | 靱 | 97E7 | 韧 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 97CC | 韌 | 97CC | 韌 | ≡ | r-trad |  | identity |
| 97CC | 韌 | 97E7 | 韧 | **→** | simp |  |  |
| **←** | trad |  |  |
| 97E7 | 韧 | 97E7 | 韧 | ≡ | 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. More information is needed to evaluate this.

The code point U+994D was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 81B3 | 膳 | 81B3 | 膳 | ≡ | r-both |  | identity |
| 81B3 | 膳 | 994D | 饍 | **→** | blocked |  |  |
| **←** | both |  |  |
| 994D | 饍 | 994D | 饍 | ≡ | r-neither |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 81B3 | 膳 | 81B3 | 膳 | ≡ | r-both |  | identity |
| 81B3 | 膳 | 994D | 饍 | ↔ | blocked |  |  |
| 994D | 饍 | 994D | 饍 | ≡ | 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’. This change needs to be justified.

The code point U+8420 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between dotAsia (table follows) and CLGR10 (where it is a singleton reflexive variant of ‘r-both’).

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

The code point U+84DA was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 84DA | 蓚 | 84DA | 蓚 | ≡ | r-neither |  | identity |
| 84DA | 蓚 | 84E8 | 蓨 | **→** | both |  |  |
| **←** | blocked |  |  |
| 84E8 | 蓨 | 84E8 | 蓨 | ≡ | r-both |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 84DA | 蓚 | 84DA | 蓚 | ≡ | r-both |  | identity |
| 84DA | 蓚 | 84E8 | 蓨 | ↔ | blocked |  |  |
| 84E8 | 蓨 | 84E8 | 蓨 | ≡ | r-both |  | identity |

Unihan does not have variant relationship for these code points, however they have common kCantonese (tiu4), and kMandarin (tiáo) values; more information is needed to evaluate this.

This variant set has one added member U+9421 (not in dotAsia.) from IICORE, J source, out of repertoire. In addition, U+9244 (in dotAsia) from IICORE HKSCS was also included in the CLGR10 (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 | 銕 | ↔ | blocked |  |  |
| 9244 | 鉄 | 9421 | 鐡 | ↔ | blocked |  |  |
| 9244 | 鉄 | 9435 | 鐵 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 9244 | 鉄 | 94C1 | 铁 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 9295 | 銕 | 9295 | 銕 | ≡ | r-trad |  | identity |
| 9295 | 銕 | 9421 | 鐡 | ↔ | blocked |  |  |
| 9295 | 銕 | 9435 | 鐵 | ↔ | blocked |  |  |
| 9295 | 銕 | 94C1 | 铁 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 9421 | 鐡 | 9421 | 鐡 | ≡ | r-neither |  | identity |
| 9421 | 鐡 | 9435 | 鐵 | **→** | trad |  |  |
| **←** | blocked |  |  |
| 9421 | 鐡 | 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 | 鉄 | 9244 | 鉄 | ≡ | r-trad |  | identity |
| 9244 | 鉄 | 9295 | 銕 | ↔ | blocked |  |  |
| 9244 | 鉄 | 9435 | 鐵 | ↔ | blocked |  |  |
| 9244 | 鉄 | 94C1 | 铁 | **→** | simp |  |  |
| **←** | trad |  |  |
| 9295 | 銕 | 9295 | 銕 | ≡ | r-trad |  | identity |
| 9295 | 銕 | 9435 | 鐵 | ↔ | blocked |  |  |
| 9295 | 銕 | 94C1 | 铁 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 9435 | 鐵 | 9435 | 鐵 | ≡ | r-trad |  | identity |
| 9435 | 鐵 | 94C1 | 铁 | **→** | simp |  |  |
| **←** | trad |  |  |
| 94C1 | 铁 | 94C1 | 铁 | ≡ | r-simp |  | identity |

Because the code point U+9421 mapping type is ‘out-of-repertoire-var’, all its mapping to other code points must be ‘blocked’.

The differences concerning U+9244 mappings are not addressed at this point (Unihan has Semantic Variant mappings between U+9244, U+9295, and U+9435). This needs further justification.

The code point U+98C7 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 98C6 | 飆 | 98C6 | 飆 | ≡ | r-trad |  | identity |
| 98C6 | 飆 | 98C7 | 飇 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 98C6 | 飆 | 98C8 | 飈 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 98C6 | 飆 | 98D9 | 飙 | **→** | simp |  |  |
| **←** | trad |  |  |
| 98C6 | 飆 | 98DA | 飚 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 98C7 | 飇 | 98C7 | 飇 | ≡ | r-neither |  | identity |
| 98C7 | 飇 | 98C8 | 飈 | ↔ | blocked |  |  |
| 98C7 | 飇 | 98D9 | 飙 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 98C7 | 飇 | 98DA | 飚 | ↔ | blocked |  |  |
| 98C8 | 飈 | 98C8 | 飈 | ≡ | r-neither |  | identity |
| 98C8 | 飈 | 98D9 | 飙 | ↔ | blocked |  |  |
| 98C8 | 飈 | 98DA | 飚 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 98D9 | 飙 | 98D9 | 飙 | ≡ | r-simp |  | identity |
| 98D9 | 飙 | 98DA | 飚 | ↔ | blocked |  |  |
| 98DA | 飚 | 98DA | 飚 | ≡ | r-simp |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 98C6 | 飆 | 98C6 | 飆 | ≡ | r-trad |  | identity |
| 98C6 | 飆 | 98C7 | 飇 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 98C6 | 飆 | 98C8 | 飈 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 98C6 | 飆 | 98D9 | 飙 | **→** | simp |  |  |
| **←** | trad |  |  |
| 98C6 | 飆 | 98DA | 飚 | **→** | blocked |  |  |
| **←** | trad |  |  |
| 98C7 | 飇 | 98C7 | 飇 | ≡ | r-neither |  | identity |
| 98C7 | 飇 | 98C8 | 飈 | ↔ | blocked |  |  |
| 98C7 | 飇 | 98D9 | 飙 | ↔ | blocked |  |  |
| 98C7 | 飇 | 98DA | 飚 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 98C8 | 飈 | 98C8 | 飈 | ≡ | r-neither |  | identity |
| 98C8 | 飈 | 98D9 | 飙 | ↔ | blocked |  |  |
| 98C8 | 飈 | 98DA | 飚 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 98D9 | 飙 | 98D9 | 飙 | ≡ | r-simp |  | identity |
| 98D9 | 飙 | 98DA | 飚 | ↔ | blocked |  |  |
| 98DA | 飚 | 98DA | 飚 | ≡ | r-simp |  | identity |

Unihan has two trad-simp variant relationship between the pairs (U+98C6, U+98D9) and (U+98C8, U+98DA), but U+98C7 is related to other code points. More information is needed to evaluate this.

The code point U+9D44 was included in CLGR10 because of its membership in IICORE HKSCS but has been assigned different types for its variant mappings between CLGR10 (1st) and dotAsia (2nd).

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

Unihan indicates that U+9D1F and U+9E31 are in a trad-simp relationship and that U+9D44 is a kDefinition variant of U+9D1F, but this does not provide a clear answer to determine which of the two LGRs is better.

## Variants sets with pre-existing code points pre-integrated with KLGR

For the following variant sets, all code points were already present in the original dotcn/dottw tables and were not the result of augmentations to sync with the dotAsia repertoire. However, the variant sets have been modified as a result of pre-integration of the KLGR repertoire.

In general, once the principle of KLGR pre-integration is accepted, most of the changes made by CLGR10 are acceptable, with the caveat that some of these changes break trad-simp mappings. Note however, that many of these changes also affect mapping types for the other characters of the mapping sets.

The following table shows a list of 146 groups of variant sets affected by that pre-integration. There is no need for a detailed analysis for most of these sets, because they are now explicitly referenced by CLGR10 in appendix D 6.2.4-441. Some of the lines correspond to multiple variant sets either as some former members transformed into singleton (yellow highlight) or moved to different set (green highlight).

Finally, some of the pre-integration entries were moved to the next section when the rationale could not be found (like having a single or no KLGR repertoire entries in the variant set).

### Table of differences from pre-integration – 146 groups of variant sets

Legend:

* Code points in red represent KLGR members,
* Code points in green represent values that CLGR identifies as KLGR members but are not part of the latest KLGR repertoire,
* Yellow-highlighted code points represent singleton reflexive variants in CLGR10,
* Green-highlighted code points represent recent HKSCS additions,
* \* represents more complex cases than just splitting the variant set as the new setting modifies mapping between remaining members, or involves multiple original variant sets
* \*\* represents hybrid cases involving pre-integration and recent HKSCS additions to CLGR, as such they are covered in separate sub-sections,
* Digits (1 to 6) in the last column indicate numbered items in the following sub-section in 5.3.2 and correspond to apparent pre-integration discrepancies because the original variant set has only one true KLGR member.

| **No** | **Variant set(s) glyphs and code points** | | | | | | | | | **N.** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 㧛39DB | 㩜3A5C | 揽63FD | 擥64E5 | 攬652C |  |  |  |  | \*\* |
| 2 | 丑4E11 | 醜919C |  |  |  |  |  |  |  |  |
| 3 | 丘4E18 | 丠4E20 | 坵5775 |  |  |  |  |  |  |  |
| 4 | 並4E26 | 併4F75 | 倂5002 | 并5E76 | 幷5E77 | 竝7ADD |  |  |  |  |
| 5 | 乾4E7E | 亁4E81 | 干5E72 | 幹5E79 | 榦69A6 | 漧6F27 |  |  |  | \* |
| 6 | 了4E86 | 暸66B8 | 瞭77AD |  |  |  |  |  |  | \* |
| 7 | 仇4EC7 | 讎8B8E | 讐8B90 | 雠96E0 |  |  |  |  |  |  |
| 8 | 仏4ECF | 佛4F5B | 彿5F7F | 髴9AF4 |  |  |  |  |  | \* |
| 9 | 仙4ED9 | 仚4EDA | 僊50CA |  |  |  |  |  |  |  |
| 10 | 仝4EDD | 同540C | 衕8855 |  |  |  |  |  |  | \* |
| 11 | 仟4EDF | 千5343 | 鱶9C76 |  |  |  |  |  |  |  |
| 12 | 仿4EFF | 倣5023 | 彷5F77 | 髣9AE3 |  |  |  |  |  |  |
| 13 | 余4F59 | 餘9918 | 馀9980 |  |  |  |  |  |  | \* |
| 14 | 侄4F84 | 姪59EA |  |  |  |  |  |  |  |  |
| 15 | 係4FC2 | 系7CFB | 繋7E4B | 繫7E6B |  |  |  |  |  | \*\* |
| 16 | 俊4FCA | 儁5101 |  |  |  |  |  |  |  |  |
| 17 | 俛4FDB | 俯4FEF | 頫982B |  |  |  |  |  |  | 1 |
| 18 | 修4FEE | 脩8129 |  |  |  |  |  |  |  |  |
| 19 | 借501F | 耤8024 | 豙8C59 |  |  |  |  |  |  | \* |
| 20 | 倦5026 | 券5238 | 劵52B5 |  |  |  |  |  |  | \* |
| 21 | 光5149 | 灮706E | 炚709A | 烡70E1 |  |  |  |  |  |  |
| 22 | 克514B | 剋524B | 尅5C05 |  |  |  |  |  |  | \* |
| 23 | 円5186 | 圆5706 | 圎570E | 圓5713 |  |  |  |  |  |  |
| 24 | 冑5191 | 胄80C4 |  |  |  |  |  |  |  |  |
| 25 | 冽51BD | 洌6D0C |  |  |  |  |  |  |  |  |
| 26 | 凄51C4 | 悽60BD | 淒6DD2 |  |  |  |  |  |  |  |
| 27 | 几51E0 | 幾5E7E |  |  |  |  |  |  |  | 2 |
| 28 | 刊520A | 刋520B | 栞681E |  |  |  |  |  |  | \* |
| 29 | 制5236 | 製88FD |  |  |  |  |  |  |  |  |
| 30 | 剳5273 | 劄5284 | 扎624E | 札672D | 箚7B9A | 紥7D25 | 紮7D2E |  |  | \* |
| 31 | 剽527D | 慓6153 |  |  |  |  |  |  |  |  |
| 32 | 办529E | 辦8FA6 | 辧8FA7 | 辨8FA8 | 釆91C6 |  |  |  |  |  |
| 33 | 勋52CB | 勛52DB | 勲52F2 | 勳52F3 |  |  |  |  |  | \* |
| 34 | 勣52E3 | 績7E3E | 绩7EE9 |  |  |  |  |  |  |  |
| 35 | 勤52E4 | 懃61C3 | 瘽763D |  |  |  |  |  |  |  |
| 36 | 勾52FE | 句53E5 |  |  |  |  |  |  |  |  |
| 37 | 匹5339 | 疋758B |  |  |  |  |  |  |  |  |
| 38 | 卜535C | 菔83D4 | 蔔8514 |  |  |  |  |  |  | \* |
| 39 | 卷5377 | 巻5DFB | 捲6372 |  |  |  |  |  |  |  |
| 40 | 历5386 | 厤53A4 | 暦66A6 | 曆66C6 | 歴6B74 | 歷6B77 |  |  |  | \* |
| 41 | 叠53E0 | 曡66E1 | 疉7589 | 疊758A | 迭8FED |  |  |  |  |  |
| 42 | 只53EA | 衹8879 | 隻96BB |  |  |  |  |  |  | \* |
| 43 | 台53F0 | 檯6AAF | 籉7C49 | 臺81FA | 颱98B1 |  |  |  |  |  |
| 44 | 叹53F9 | 嘆5606 | 歎6B4E |  |  |  |  |  |  |  |
| 45 | 合5408 | 敆6546 | 閤95A4 |  |  |  |  |  |  |  |
| 46 | 向5411 | 嚮56AE | 曏66CF |  |  |  |  |  |  | \* |
| 47 | 呵5475 | 嗬55EC | 訶8A36 | 诃8BC3 |  |  |  |  |  |  |
| 48 | 咀5480 | 嘴5634 |  |  |  |  |  |  |  |  |
| 49 | 咸54B8 | 鹹9E79 |  |  |  |  |  |  |  |  |
| 50 | 咽54BD | 嚥56A5 |  |  |  |  |  |  |  |  |
| 51 | 哲54F2 | 喆5586 |  |  |  |  |  |  |  |  |
| 52 | 唇5507 | 脣8123 |  |  |  |  |  |  |  |  |
| 53 | 喜559C | 憙6199 |  |  |  |  |  |  |  |  |
| 54 | 喰55B0 | 飡98E1 | 餐9910 |  |  |  |  |  |  | \* |
| 55 | 囘56D8 | 回56DE | 囬56EC | 廻5EFB | 廽5EFD | 迴8FF4 | 逥9025 |  |  | \* |
| 56 | 圣5723 | 垩57A9 | 堊580A | 聖8056 |  |  |  |  |  | \* |
| 57 | 圭572D | 珪73EA |  |  |  |  |  |  |  |  |
| 58 | 圻573B | 垠57A0 |  |  |  |  |  |  |  |  |
| 59 | 坂5742 | 阪962A |  |  |  |  |  |  |  |  |
| 60 | 坐5750 | 座5EA7 |  |  |  |  |  |  |  |  |
| 61 | 埙57D9 | 塤5864 | 壎58CE |  |  |  |  |  |  |  |
| 62 | 埼57FC | 崎5D0E | 碕7895 |  |  |  |  |  |  | \* |
| 63 | 堈5808 | 缸7F38 |  |  |  |  |  |  |  |  |
| 64 | 塙5859 | 确786E | 確78BA | 碻78BB | 礭792D |  |  |  |  | \* |
| 65 | 复590D | 復5FA9 | 複8907 | 覆8986 |  |  |  |  |  | \* |
| 66 | 奈5948 | 柰67F0 |  |  |  |  |  |  |  |  |
| 67 | 姜59DC | 薑8591 |  |  |  |  |  |  |  |  |
| 68 | 娘5A18 | 嬢5B22 | 孃5B43 |  |  |  |  |  |  | \* |
| 69 | 孚5B5A | 孵5B75 |  |  |  |  |  |  |  |  |
| 70 | 宁5B81 | 寍5BCD | 寕5BD5 | 寗5BD7 | 寜5BDC | 寧5BE7 | 甯752F |  |  | \* |
| 71 | 尿5C3F | 溺6EBA |  |  |  |  |  |  |  |  |
| 72 | 岭5CAD | 岺5CBA | 嶺5DBA |  |  |  |  |  |  | \* |
| 73 | 巨5DE8 | 鉅9245 | 钜949C |  |  |  |  |  |  |  |
| 74 | 希5E0C | 稀7A00 |  |  |  |  |  |  |  |  |
| 75 | 年5E74 | 秊79CA |  |  |  |  |  |  |  |  |
| 76 | 庵5EB5 | 菴83F4 |  |  |  |  |  |  |  |  |
| 77 | 彫5F6B | 琱7431 | 雕96D5 | 鵰9D70 |  |  |  |  |  | \* |
| 78 | 征5F81 | 徴5FB4 | 徵5FB5 |  |  |  |  |  |  | \* |
| 79 | 径5F84 | 徑5F91 | 迳8FF3 | 逕9015 |  |  |  |  |  |  |
| 80 | 御5FA1 | 禦79A6 |  |  |  |  |  |  |  | \* |
| 81 | 徧5FA7 | 遍904D |  |  |  |  |  |  |  | 3 |
| 82 | 徨5FA8 | 遑9051 |  |  |  |  |  |  |  |  |
| 83 | 徳5FB3 | 德5FB7 | 悳60B3 | 惪60EA |  |  |  |  |  | \* |
| 84 | 怜601C | 憐6190 |  |  |  |  |  |  |  |  |
| 85 | 总603B | 惣60E3 | 揔63D4 | 捴6374 | 搃6403 | 摠6460 | 総7DCF | 縂7E02 | 總7E3D | \*\* |
| 86 | 恂6042 | 悛609B |  |  |  |  |  |  |  |  |
| 87 | 恵6075 | 惠60E0 | 憓6193 |  |  |  |  |  |  |  |
| 88 | 愈6108 | 瘉7609 | 癒7652 |  |  |  |  |  |  | \* |
| 89 | 愿613F | 願9858 |  |  |  |  |  |  |  |  |
| 90 | 慄6144 | 栗6817 |  |  |  |  |  |  |  | \* |
| 91 | 慇6147 | 殷6BB7 |  |  |  |  |  |  |  |  |
| 92 | 慧6167 | 譿8B7F |  |  |  |  |  |  |  |  |
| 93 | 慾617E | 欲6B32 |  |  |  |  |  |  |  |  |
| 94 | 懽61FD | 欢6B22 | 歓6B53 | 歡6B61 | 讙8B99 | 驩9A69 |  |  |  |  |
| 95 | 扺627A | 抵62B5 | 牴7274 | 羝7F9D | 觝89DD |  |  |  |  | 4 |
| 96 | 拏62CF | 拿62FF |  |  |  |  |  |  |  |  |
| 97 | 拑62D1 | 箝7B9D | 鉗9257 | 钳94B3 |  |  |  |  |  | \* |
| 98 | 挽633D | 輓8F13 |  |  |  |  |  |  |  |  |
| 99 | 捨6368 | 舍820D | 舎820E |  |  |  |  |  |  | \* |
| 99 | 捻637B | 撚649A |  |  |  |  |  |  |  |  |
| 100 | 捿637F | 栖6816 | 棲68F2 |  |  |  |  |  |  | \*\* |
| 101 | 斗6597 | 闘95D8 | 鬥9B25 | 鬦9B26 | 鬪9B2A | 鬬9B2C | 鬭9B2D |  |  | \* |
| 102 | 斤65A4 | 筋7B4B | 觔89D4 |  |  |  |  |  |  | \* |
| 103 | 昞661E | 昺663A | 炳70B3 |  |  |  |  |  |  |  |
| 104 | 映6620 | 暎668E |  |  |  |  |  |  |  |  |
| 105 | 是662F | 昰6630 |  |  |  |  |  |  |  |  |
| 106 | 晃6643 | 晄6644 |  |  |  |  |  |  |  |  |
| 107 | 晧6667 | 暠66A0 | 皓7693 | 皜769C |  |  |  |  |  | \* |
| 108 | 晻667B | 暗6697 | 闇95C7 |  |  |  |  |  |  |  |
| 109 | 曌66CC | 炤70A4 | 照7167 |  |  |  |  |  |  |  |
| 110 | 曲66F2 | 麯9EAF | 麴9EB4 | 麹9EB9 |  |  |  |  |  | \*\* |
| 111 | 曹66F9 | 曺66FA |  |  |  |  |  |  |  |  |
| 112 | 朞671E | 期671F |  |  |  |  |  |  |  |  |
| 113 | 朮672E | 术672F | 術8853 |  |  |  |  |  |  | \* |
| 114 | 朴6734 | 樸6A38 |  |  |  |  |  |  |  | \* |
| 115 | 村6751 | 邨90A8 |  |  |  |  |  |  |  |  |
| 116 | 枏678F | 楠6960 |  |  |  |  |  |  |  |  |
| 117 | 柒67D2 | 桼687C | 漆6F06 |  |  |  |  |  |  | \* |
| 118 | 柩67E9 | 柾67FE |  |  |  |  |  |  |  |  |
| 119 | 桁6841 | 航822A |  |  |  |  |  |  |  |  |
| 120 | 桌684C | 棹68F9 | 櫂6AC2 |  |  |  |  |  |  |  |
| 121 | 櫱6AF1 | 蘖8616 |  |  |  |  |  |  |  | 5 |
| 122 | 欝6B1D | 郁90C1 | 鬰9B30 | 鬱9B31 |  |  |  |  |  | \* |
| 123 | 氾6C3E | 汎6C4E | 泛6CDB |  |  |  |  |  |  | \* |
| 124 | 沾6CBE | 霑9711 |  |  |  |  |  |  |  |  |
| 125 | 浜6D5C | 滨6EE8 | 濱6FF1 |  |  |  |  |  |  | \* |
| 126 | 浣6D63 | 澣6FA3 |  |  |  |  |  |  |  |  |
| 127 | 淋6DCB | 痲75F2 | 痳75F3 | 蔴8534 | 麻9EBB |  |  |  |  | \* |
| 128 | 淡6DE1 | 澹6FB9 |  |  |  |  |  |  |  |  |
| 129 | 漾6F3E | 瀁7001 |  |  |  |  |  |  |  |  |
| 130 | 炼70BC | 煉7149 | 錬932C | 鍊934A |  |  |  |  |  | \* |
| 131 | 狯72EF | 猾733E | 獪736A |  |  |  |  |  |  |  |
| 132 | 獲7372 | 穫7A6B | 获83B7 |  |  |  |  |  |  | \* |
| 133 | 玟739F | 珉73C9 | 琝741D |  |  |  |  |  |  | \* |
| 134 | 着7740 | 著8457 |  |  |  |  |  |  |  | \* |
| 135 | 祕7955 | 秘79D8 |  |  |  |  |  |  |  |  |
| 136 | 穀7A40 | 谷8C37 |  |  |  |  |  |  |  |  |
| 137 | 緻7DFB | 致81F4 |  |  |  |  |  |  |  | \* |
| 138 | 胡80E1 | 衚885A | 鬍9B0D |  |  |  |  |  |  | 6 |
| 139 | 舡8221 | 舩8229 | 船8239 |  |  |  |  |  |  | \* |
| 140 | 芸82B8 | 蕓8553 |  |  |  |  |  |  |  | \* |
| 141 | 苽82FD | 菇83C7 | 菰83F0 |  |  |  |  |  |  |  |
| 142 | 裏88CF | 裡88E1 | 里91CC |  |  |  |  |  |  | \* |
| 143 | 辟8F9F | 避907F | 闢95E2 |  |  |  |  |  |  |  |
| 144 | 酢9162 | 醋918B |  |  |  |  |  |  |  |  |
| 145 | 錘9318 | 鎚939A | 锤9524 |  |  |  |  |  |  |  |
| 146 | 須9808 | 须987B | 鬚9B1A |  |  |  |  |  |  |  |

\*\*

The following cases are a hybrid of pre-integration and recently added HKSCS members (not in dotAsia); the pre-integration does not seem problematic, but further discussion concerning the additions is provided in section 5.1.4.

Case #1 is a mix of CLGR/KLGR pre-integration (affecting U+64E5 and U+652C) and one added code point: U+3A5C. The addition is examined separately in 5.1.4.4.

Case #15 is also a mix of CLGR/KLGR pre-integration (affecting U+4FC2, U+7CFB, and U+7E6B) and one added code point: U+7E4B. The addition is examined separately in 5.1.4.2.

Case #110 is also a mix of CLGR/KLGR pre-integration (affecting U+66F2 and U+9EB4) and one added code point: U+9EB9. The addition is examined separately in see 5.1.4.8

The following cases are a hybrid of pre-integration and recently added HKSCS members through dotAsia; the pre-integration does not seem problematic, but further discussion concerning the additions is provided n section 5.2.2.

Case #85 is also a mix of CLGR/KLGR pre-integration (affecting U+6460 and U+7E3D) and a code point added through dotAsia (U+60E3) treated differently. The difference concerning U+60E3 is examined separately in 5.2.2.9.

Case #100 is also a mix of CLGR/KLGR pre-integration (affecting U+637F, U+6816, and U+68F2) and a code point added through dotAsia (U+637F) treated differently. The difference concerning U+637F is examined separately in 5.2.2.12.

### Detailed list of KLGR pre-integration discrepancies – 6 variant sets

All six sets in this section are referenced by CLGR10 in appendix D 6.2.4-441 as showing at least two code points claimed as being part of the KLGR repertoire, which justifies the modification as a result of pre-integration. However, on further review the IP found that the actual KLGR repertoire only contains one of these code points, the other is missing from the KLGR repertoire. Upon querying the KGP and CGP, the IP has learned that the cause for that may be the reduction of the Hanja set defined in KLGR.

At this point, these six variant sets modifications are problematic, because there no longer is an impact on KLGR. The IP requests the CGP either to withdraw these changes, reverting to established practice, or to provide documentation of an independent rationale that supports a preference for the new variant sets. In either case, the documentation by CGP needs to be adjusted to reflect the actual members of the KLGR repertoire.

In CLGR10 (1st table), U+4FDB and U+982B have a variant relationship, U+4FEF is a singleton reflexive variant of type ‘r-both’. In dotAsia (2nd table), U+4FDB, U+4FEF, and U+982B do have a variant relationship.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 4FDB | 俛 | 4FDB | 俛 | ≡ | r-both |  | identity |
| 4FDB | 俛 | 982B | 頫 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 982B | 頫 | 982B | 頫 | ≡ | r-trad |  | identity |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 4FDB | 俛 | 4FDB | 俛 | ≡ | r-trad |  | identity |
| 4FDB | 俛 | 4FEF | 俯 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 4FDB | 俛 | 982B | 頫 | ↔ | blocked |  |  |
| 4FEF | 俯 | 4FEF | 俯 | ≡ | r-both |  | identity |
| 4FEF | 俯 | 982B | 頫 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 982B | 頫 | 982B | 頫 | ≡ | r-trad |  | identity |

Unihan has variant relationship for U+4FEF and U+982B, U+9980, (but not U+4FDB); only U+4FEF is part of KLGR (appendix D 6.2.4-441 says that **U+4FDB** is part of KGP, but this is not the case in latest KLGR).

In this case, both variant sets seem problematic. According to Unihan, only U+4FEF and U+982B should be member of the variant set. In the latest update to CLGR, U+ 4FDB is no longer listed as a member of the KLGR repertoire (reference 500).

In dotAsia (table), U+51E0 and U+5E7E have a variant relationship. In CLGR10, U+51E0 and U+5E7E are singleton reflexive variant of type ‘r-both’.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 51E0 | 几 | 51E0 | 几 | ≡ | r-both |  | identity |
| 51E0 | 几 | 5E7E | 幾 | **→** | trad |  |  |
| **←** | simp |  |  |
| 5E7E | 幾 | 5E7E | 幾 | ≡ | r-trad |  | identity |

Unihan has ‘Trad-Simp’ variant relationship between both code points. Only U+5E7E is part of KLGR (appendix D 6.2.4-441 says that **U+51E0** is part of KGP, but this is not the case in latest KLGR). It is not clear why this variant set was removed. In the latest update to CLGR, U+51E0 is no longer listed as a member of the KLGR repertoire (reference 500).

In dotAsia U+5FA7 and U+904D (table follows) have a variant relationship. In CLGR10, both code points are singleton reflexive variant of type ‘r-both’.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 5FA7 | 徧 | 5FA7 | 徧 | ≡ | r-neither |  | identity |
| 5FA7 | 徧 | 904D | 遍 | **→** | both |  |  |
| **←** | blocked |  |  |
| 904D | 遍 | 904D | 遍 | ≡ | r-both |  | identity |

Unihan has a variant relationship between these code points. Only the code points U+904D is part of KLGR (appendix D 6.2.4-441 says that **U+5FA7** is part of KGP, but this is not the case in latest KLGR). It is not clear why this variant set was removed. In the latest update to CLGR, U+5FA7 is no longer listed as a member of the KLGR repertoire (reference 500).

In dotAsia U+627A, U+62B5, U+7274, U+7F9D, and U+89DD (table follows) have a variant relationship. In CLGR10, U+89DD is a singleton reflexive variant of type ‘r-both’.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 627A | 扺 | 627A | 扺 | ≡ | r-both |  | identity |
| 627A | 扺 | 62B5 | 抵 | ↔ | blocked |  |  |
| 627A | 扺 | 7274 | 牴 | ↔ | blocked |  |  |
| 627A | 扺 | 7F9D | 羝 | ↔ | blocked |  |  |
| 627A | 扺 | 89DD | 觝 | ↔ | blocked |  |  |
| 62B5 | 抵 | 62B5 | 抵 | ≡ | r-both |  | identity |
| 62B5 | 抵 | 7274 | 牴 | ↔ | blocked |  |  |
| 62B5 | 抵 | 7F9D | 羝 | ↔ | blocked |  |  |
| 62B5 | 抵 | 89DD | 觝 | **→** | trad |  |  |
| **←** | simp |  |  |
| 7274 | 牴 | 7274 | 牴 | ≡ | r-both |  | identity |
| 7274 | 牴 | 7F9D | 羝 | ↔ | blocked |  |  |
| 7274 | 牴 | 89DD | 觝 | ↔ | blocked |  |  |
| 7F9D | 羝 | 7F9D | 羝 | ≡ | r-both |  | identity |
| 7F9D | 羝 | 89DD | 觝 | ↔ | blocked |  |  |
| 89DD | 觝 | 89DD | 觝 | ≡ | r-trad |  | identity |

Unihan has a variant relationship between sets (U+627A, U+62B5) and (U+72F4, U+7274, U+7F9D, and U+89DD). Only the code points U+62B5 is part of KLGR (appendix D 6.2.4-441 says that **U+89DD** is part of KGP, but this is not the case in latest KLGR). It is not clear why this variant set was removed. In the latest update to CLGR, U+ 89DD is no longer listed as a member of the KLGR repertoire (reference 500).

In dotAsia U+6AF1 and U+8616 (table follows) have a variant relationship. In CLGR10, both are singleton reflexive variant of type ‘r-both’.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 6AF1 | 櫱 | 6AF1 | 櫱 | ≡ | r-trad |  | identity |
| 6AF1 | 櫱 | 8616 | 蘖 | **→** | simp |  |  |
| **←** | blocked |  |  |
| 8616 | 蘖 | 8616 | 蘖 | ≡ | r-both |  | identity |

Unihan has a trad-simp variant relationship between these code points. Only the code point U+8616 is part of KLGR (appendix D 6.2.4-441 says that **U+6AF1** is part of KGP, but this is not the case in latest KLGR). It is not clear why this variant set was removed. In the latest update to CLGR, U+ 6AF1 is no longer listed as a member of the KLGR repertoire (reference 500).

In dotAsia U+80E1, U+885A and U+9B0D (table follows) have a variant relationship. In CLGR10, U+9B0D is a singleton reflexive variant of type ‘r-both’, all other mappings unchanged.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 80E1 | 胡 | 80E1 | 胡 | ≡ | r-both |  | identity |
| 80E1 | 胡 | 885A | 衚 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 80E1 | 胡 | 9B0D | 鬍 | **→** | trad |  |  |
| **←** | simp |  |  |
| 885A | 衚 | 885A | 衚 | ≡ | r-trad |  | identity |
| 885A | 衚 | 9B0D | 鬍 | ↔ | blocked |  |  |
| 9B0D | 鬍 | 9B0D | 鬍 | ≡ | r-trad |  | identity |

Unihan has a simp-trad variant relationship between U+80E1 (simp) and U+885A or U+9B0D (both trad), and a zVariant with U+659C. Only U+80E1 part of KLGR (appendix D 6.2.4-441 says that **U+9B0D** is part of KGP, but this is not the case in latest KLGR). This split breaks a trad-simp mapping. It is not clear why this variant mapping was removed. In the latest update to CLGR, U+ 9B0D is no longer listed as a member of the KLGR repertoire (reference 500).

## Variant sets with pre-existing code points but different variant types – 3 sets

For the following variant sets, all code points were already present in the original dotcn/dottw tables and were not the result of augmentations to synchronize with the dotAsia repertoire. However, the assigned variant types in CLGR10 do not agree with the earlier tables. The variant sets listed below show deviation between CLGR10 and the original sources.

Note that in addition to these 3 variant sets, there are 2 entries from a previous section corresponding to the additions of code points U+61F4 and U+74A2 (section 5.1.4) that need the same documentation because they also contain differences in variant types for some pre-existing code points (in addition to other differences).

As with the earlier versions of CLGR, the IP has not yet evaluated the merits of the changes listed below, and before doing so, would like to request the CGP to provide some background information, rationale, references etc. that explain these differences.

The 2 code points U+5B0E and U+5B14 form their own variant set in dotAsia. In CLGR10, these two code points are singletons with reflexive variant of type ‘r-both’.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 5B0E | 嬎 | 5B0E | 嬎 | ≡ | r-both |  | identity |
| 5B0E | 嬎 | 5B14 | 嬔 | ↔ | blocked |  |  |
| 5B14 | 嬔 | 5B14 | 嬔 | ≡ | r-both |  | identity |

Unihan does not provide any clues concerning this (no variant set for any of these code points). More information is needed to evaluate this. None of the code points is part of KLGR. The glyph differences appear very minuscule which may or may not have been a factor in the earlier treatment as blocked variant. If this is a case of adding a missing variant mapping, it should be explicitly documented as such.

In these tables, the types of the mappings between U+67A3 and U+6806 or U+68D7 are reversed between CLGR10 (1st) and dotAsia (2nd).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 67A3 | 枣 | 67A3 | 枣 | ≡ | r-simp |  | identity |
| 67A3 | 枣 | 6806 | 栆 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 67A3 | 枣 | 68D7 | 棗 | **→** | trad |  |  |
| **←** | simp |  |  |
| 6806 | 栆 | 6806 | 栆 | ≡ | r-trad |  | identity |
| 6806 | 栆 | 68D7 | 棗 | ↔ | blocked |  |  |
| 68D7 | 棗 | 68D7 | 棗 | ≡ | r-trad |  | identity |
| **Source** | **Glyph** | **Target** | **Glyph** |  | **Type(s)** | **Ref** | **Comment** |
| 67A3 | 枣 | 67A3 | 枣 | ≡ | r-simp |  | identity |
| 67A3 | 枣 | 6806 | 栆 | **→** | trad |  |  |
| **←** | simp |  |  |
| 67A3 | 枣 | 68D7 | 棗 | **→** | blocked |  |  |
| **←** | simp |  |  |
| 6806 | 栆 | 6806 | 栆 | ≡ | r-trad |  | identity |
| 6806 | 栆 | 68D7 | 棗 | ↔ | blocked |  |  |
| 68D7 | 棗 | 68D7 | 棗 | ≡ | r-trad |  | identity |

Unihan confirms the CLGR10 option (U+68D7 kTraditionalVariant for U+67A3). If this is a case of correcting a mistaken variant mapping, it should be explicitly documented as such.

In CLGR10 U+68C5 maps to U+67C4 with a type of ‘both’ as shown in the following table. This is different from both .cn and dotAsia where U+68C5 maps to itself as ‘r-both’ type and where there is no variant relation between these code points (i.e. they are singletons).

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

Unihan defines kSemanticVariant mapping between U+67C4 and U+68C5, which could support the CLGR10 mapping. If this is a case of adding a missing variant mapping, it should be explicitly documented as such.

# Mechanism for reducing multiple allocatable labels

The proposed mechanism in Section 6.3 of [Proposal] is a feasible approach, at the expense of additional complexity in the WLE rules.

1. The actions defined will not assign an allocatable status to the original label, which is presumably not the intention of the CGP. The following modified actions may result in the intended behavior:

<action disp="invalid" any-variant="out-of-repertoire-var" comment="action for imported variant" />

<action disp="blocked" any-variant="blocked" comment="default action for blocked variant"/>

<action disp="allocatable" only-variants="simp r-simp both r-both r-both-mt" comment="all simplified label" />

<action disp="allocatable" only-variants="trad r-trad both r-both r-both-ms trad-1" comment="all traditional label category one"/>

<action disp="allocatable" only-variants="trad r-trad both r-both r-both-ms trad-2" comment="all traditional label category two"/>

<action disp="blocked" any-variant="simp trad both trad-m trad-1 trad-2" comment="block any other mixed labels" />

<action disp="allocatable" all-variants="r-neither r-trad r-simp r-both r-both-mt r-both-ms" comment="original label" />

<action disp="valid" comment="catch all (default action)"/>

2. In order to allow the IP to evaluate the effects of this mechanism, the newly introduced types should be used in the variant sets in the LGR XML.

3. In the proposal text, the examples given do not correspond directly to each of the categories of multiple allocatable mappings. It would greatly help the clarity of the proposal if the examples, along with the assigned types, correspond to the categories.

4. Since “r-simp” and “r-both-mt” are functionally equivalent, as are “r-trad” instead of “r-both-ms”, consider reusing the existing types rather than introducing new types.

1. A ‘ref’ attribute can also be used to reference a more detailed discussion of some variants in the main proposal or one of its appendices. [↑](#footnote-ref-1)