

Title: Proposal to Add “kSEAL_AdditionalSrc” Property to SealSources.txt

Source: suzuki toshiya (mpsuzuki@hiroshima-u.ac.jp)

Date: 2026-09-28

Action: Discussion by CJK Experts

References:

WG2 N4834 https://unicode.org/wg2/docs/n4843Response%20to%20WG2N4834_TCAandCHINA.pdf

WG2 N4844 https://unicode.org/wg2/docs/n4844_CommentsonencodingShuowenSmallSea_SelenalandEiso.pdf

WG2 N4973 https://unicode.org/wg2/docs/n4973_comments%20on%20encoding%20Shuowen%20Small%20Seal.pdf

WG2 N5089 https://unicode.org/wg2/docs/n5089_Shuowen%20Seal%20Informal%20Meeting%20Report.pdf

WG2 N5230 <https://unicode.org/wg2/docs/n5230-TCA%20Feedback%20to%20N5209.pdf>

WG2 N5211 <https://unicode.org/wg2/docs/n5211-14-columnchart-20171206.pdf>

Abstract

There are some glyph names that are supposed to exist in source references but do not appear in the code chart and SealSources.txt, because the duplicated encodings of the same shape glyphs are avoided, as discussed in WG2 N4834, N4844, N4973, N5089, and N5209. However, there is no information which glyphs are omitted, or which codepoints can render them in the presentations of the source data. This submission proposes adding new property to SealSources.txt to give such information.

1. Omitted Seal Glyphs in Unicode 18.0.0

Some users may try to produce digital texts, or the Seal glyph list of Shuowen Jiezi (SWJZ) by Seal characters coded in Unicode 18.0.0, they would find some glyph names are omitted in the code chart and SealSources.txt. Below is the list of omitted glyph names, the order of the source references follows the order in the code chart (THX, CCZ, QJZ and DYC):

- 17 omitted glyph names for THX:
 - TH-00881, TH-00899, TH-01029, TH-01790, TH-01791, TH-02063,
 - TH-02078, TH-02259, TH-03407, TH-04191, TH-04307, TH-05054,
 - TH-05346, TH-06060, TH-07327, TH-08352, TH-09707
- 18 omitted glyph names for CCZ:
 - C-00914, C-00932, C-01062, C-01855, C-01856, C-02141, C-02157,
 - C-02340, C-03518, C-04311, C-04439, C-05222, C-05517, C-06138,
 - C-06277, C-07601, C-08670, C-10070
- 21 omitted glyph names for QJZ:
 - K-00729, K-00887, K-00906, K-01035, K-02068, K-02083, K-02266,
 - K-03416, K-03818, K-04276, K-04314, K-05061, K-05355, K-06094,
 - K-07013, K-07050, K-07254, K-07341, K-08365, K-09722, K-10163

- 23 omitted glyph names for DYC:

D-00636, D-00637, D-00876, D-00894, D-01023, D-02046, D-02061,
D-02244, D-03397, D-03795, D-04175, D-04292, D-05039, D-05181,
D-05334, D-06050, D-06991, D-07315, D-07635, D-08341, D-08865,
D-09705, D-10147

A demo tool to list the omitted glyph names from SealSources.txt is available at:

https://mpsuzuki.github.io/kSEAL_DUP/

2. Issues Related to the Omitted Glyphs

Although these omissions were discussed and agreed by the Small Seal experts, the discussion record is not summarized in any part of the standard. The users may wonder about the gap of the glyph name sequences, and how to render the omitted glyphs. Even if they know about the duplicated entries, it would not be so easy to spot which codepoint should be used to render the omitted glyph.

3. Proposal for New Property

To clarify the omitted glyphs and their substitutions, there might be some options on how to add such information into SealSources.txt. Table 1 describes the proposed kSEAL_AdditionalSrc property in UAX#60 format. The complete list of the proposed properties is shown on page 4, and the detailed rationales for each codepoint are described in the Appendix.

Table 1: Definition of the proposed kSEAL_AdditionalSrc

Property	kSEAL_AdditionalSrc
Status	Provisional
Category	Other data
Introduced	TBD
Delimiter	space
Syntax	([CDK] TH)-\d{5}
Default	N/A
Description	One or more kSEAL_THXSrc, kSEAL_CCZSrc, kSEAL_QJZSrc, or kSEAL_DYCSrc property values that correspond to source references that were omitted due to duplicate glyphs in a source. The order of the source references follows the order in the code charts.

There might be other options to add the information on the glyphs omitted from the code chart.

- From the viewpoint of the users, changing `kSEAL_THXSrc`, `kSEAL_CCZSrc`, etc. to multi-value properties (like the `kSEAL_Rad` property) would be the most straightforward solution. The first value would appear in the code chart, but the rest values do not.
- Also, there might be another idea to change the syntax of these properties to the concatenated syntax, like `kTGT_MergedSrc` of the Tangut script. For example, using the value “TH-00888-06060” to mean “TH-00888 and TH-06060”. The first deduced value would appear in the code chart, but the rest do not. This idea would be still user-friendly too, while the extended property is still single value property.

However, considering that the existing source properties are normative and might be important to maintain the consistency between the code chart and UCD data, changing the existing property design might have some impact with the code chart production workflow. As the first proposal, this document proposes adding new provisional property to minimize the impact.

Also, the expected number of the UCS codepoints with this property would be less than 50, thus it can be a considerable option to have this attribute for each source, like, `kSEAL_THXAdditionalSrc`, `kSEAL_CCZAdditionalSrc`, etc.

Acknowledgement

This proposal contains the outcomes of the research supported by JSPS KAKENHI Grant Number JP16K004600A, JP19K12716, and JP22K12719. I would like to thank Jiang Kushim (for his suggestion of the multi-value property), Eiso Chan (for his suggestion of the concatenated syntax), and Ken Lunde for valuable discussion in drafting. This proposal would not have been possible without their support.

```

# kSEAL_AdditionalSrc
#
# UAX #60 Additions:
#
# Property:      kSEAL_AdditionalSrc
# Status:        Provisional
# Category:      Other data
# Introduced:    TBD
# Delimiter:     space
# Syntax:        ([CDK]|TH)-\d{5}
# Default:       N/A
# Description:   One or more kSEAL_THXSrc, kSEAL_CCZSrc, kSEAL_QJZSrc, or
kSEAL_DYCSrc property values that correspond to source references that
were omitted due to duplicate glyphs in a source. The order of the source
references follows the order in the code charts.
#
# 30 New SealSources.txt Records:
#
U+3D1B6 kSEAL_AdditionalSrc      K-00729
U+3D228 kSEAL_AdditionalSrc      K-07013 D-06991
U+3D3AB kSEAL_AdditionalSrc      TH-06060 C-06277 K-06094 D-06050
U+3D3DD kSEAL_AdditionalSrc      TH-02078 C-02157 K-02083 D-02061
U+3D406 kSEAL_AdditionalSrc      TH-03407 C-03518 K-03416 D-03397
U+3D421 kSEAL_AdditionalSrc      TH-08352 C-08670 K-08365 D-08341
U+3D43C kSEAL_AdditionalSrc      D-07635
U+3D626 kSEAL_AdditionalSrc      C-06138
U+3D685 kSEAL_AdditionalSrc      D-05181
U+3D72D kSEAL_AdditionalSrc      TH-01791 C-01856
U+3D72E kSEAL_AdditionalSrc      TH-01790 C-01855
U+3DB65 kSEAL_AdditionalSrc      TH-04307 C-04439 K-04314 D-04292
U+3DC96 kSEAL_AdditionalSrc      TH-02259 C-02340 K-02266 D-02244
U+3DE12 kSEAL_AdditionalSrc      TH-07327 C-07601 K-07341 D-07315
U+3DE72 kSEAL_AdditionalSrc      D-00636
U+3DE73 kSEAL_AdditionalSrc      D-00637
U+3E003 kSEAL_AdditionalSrc      TH-04191 C-04311 D-04175
U+3E0B3 kSEAL_AdditionalSrc      K-10163 D-10147
U+3E923 kSEAL_AdditionalSrc      TH-01029 C-01062 K-01035 D-01023
U+3ED9C kSEAL_AdditionalSrc      K-07050
U+3EE4E kSEAL_AdditionalSrc      TH-00899 C-00932 K-00906 D-00894
U+3EE9D kSEAL_AdditionalSrc      K-07254
U+3EFE3 kSEAL_AdditionalSrc      TH-00881 C-00914 K-00887 D-00876
U+3F2C6 kSEAL_AdditionalSrc      K-04276
U+3F322 kSEAL_AdditionalSrc      TH-02063 C-02141 K-02068 D-02046
U+3F38D kSEAL_AdditionalSrc      K-03818 D-03795
U+3F4EA kSEAL_AdditionalSrc      D-08865
U+3F792 kSEAL_AdditionalSrc      TH-09707 C-10070 K-09722 D-09705
U+3FADF kSEAL_AdditionalSrc      TH-05346 C-05517 K-05355 D-05334
U+3FB61 kSEAL_AdditionalSrc      TH-05054 C-05222 K-05061 D-05039
# EOF

```

Appendix A: Background of Each Property

In this appendix, the relationship of the omitted glyphs and standardized characters are described. There are 3 categories why there is a gap in the glyph name sequences:

- **Row omissions for universal duplicates**: all major editions (DaXu Ben, XiaoXu Ben, and Duan Zhu Ben) contain the occurrences, and all instances are duplicated, thus whole row is omitted.
- **Row omissions for edition-specific duplicates**: some but not all editions (DaXu Ben only, XiaoXu Ben + Duan Zhu Ben only, etc) contain the occurrences, but all instances are duplicated whenever they are found, thus whole row is omitted.
- **Cell omissions for edition-specific duplicates**: some editions have distinguishable instances, but others have duplicated instances. Only the duplicated instances are omitted, thus the blank cells are left in the row.

The sections below describe each category.

- A.1 Terminology: Branch and Edition
- A.2 Relevant Columns of WG2 N5211
- A.3 Row Omissions for Universal Duplicates
- A.4 Row Omissions for Edition-Specific Duplicates
- A.5 Cell Omissions for Edition-Specific Duplicates

A.1. Terminology: Branch and Edition

To describe similarities and differences among the extant texts, this document uses two terms: branch and edition. Figure 1 shows the outline of the relationship.

A printed version issued by a specific publisher is referred to as an “edition”. In most cases, a published SWJZ text is presented as a reproduction of an authoritative source rather than a compilation from multiple sources.

If the principal source is identified as the text restored by Xu Xuan, the edition is regarded as an instance of the “DaXu Ben” (大徐本) branch. If the principal source is identified as the text restored and annotated by Xu Kai, the edition is regarded as an instance of the “XiaoXu Ben” (小徐本) branch. The text restored and annotated by Duan Yucai is referred to as the “DuanZhu Ben” (段注本).

These broad categories are sometimes referred to as “DaXu versions” and “XiaoXu versions”. For greater precision, however, this document uses the terms “DaXu branch” (大徐系統) and “XiaoXu” branch (小徐系統).

The Jiguge edition, published by the Jiguge Library between the late Ming and early Qing periods, underwent significant revision between its earlier printing (“毛 4” column in Figure 5) and later printing (“毛 5” column in Figure 5). When a distinction is required, these are referred to as the Jiguge

early print and the Jiguge later print. Otherwise, both are collectively referred to as the “Jiguge edition”. The Jiguge edition is regarded as an instance of the “DaXu branch” in broad sense, but it is based on “Shuowen Jiezi Wuyun Yunpu” (說文解字五音韻譜), and introduced some entries from the XiaoXu branch. Therefore, it is not an instance of the “DaXu” branch in strict sense.

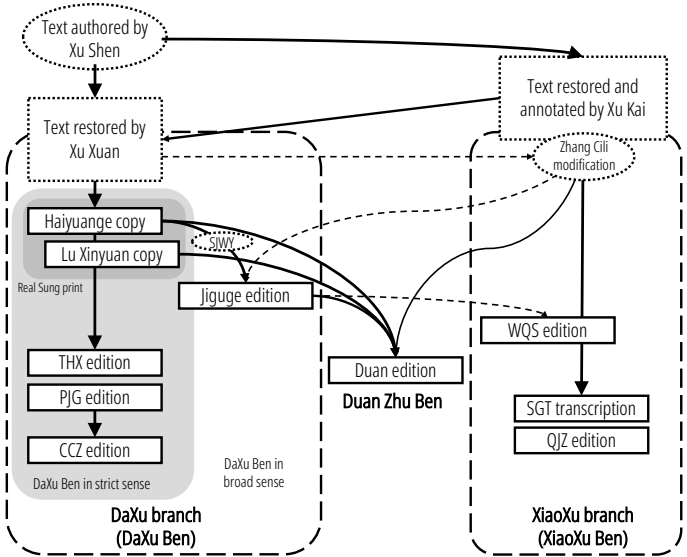


Figure 1: “Branch” and “Edition” in this document, and their relationship

A.1.1. How Omissions Appear in the Code Chart

The appearances of the omissions can be classified into 3 categories as below:

① Row omissions of universal duplicates

Figure 2 shows the case that the occurrence (the second “U”) is found in all editions (L, M, N), and all instance is duplicated and the whole row is omitted. There would be a glyph name gap between the previous (“X”) and next rows (“Y”), for all editions.

With Duplicates				Standardized		
Src L	Src M	Src N		Src L	Src M	Src N
T	T	T	→	T	T	T
1234	1232	1201		1234	1232	1201
U	U	U	→	U	U	U
1235	1233	1202		1235	1233	1202
...	...	...		...	...	...
X	X	X	→	X	X	X
1301	1402	1250		1301	1402	1250
U	U	U	↘	Y	Y	Y
1302	1403	1251		1303	1404	1252
Y	Y	Y	↗	...	...	...
1303	1404	1252				
...	...	...				

Figure 2: Row omissions of universal entries

② Row omissions of edition-dependent duplicates

Figure 3 shows the case where the occurrence (the second “U”) is found in some but not all editions, but the instance is duplicated whenever it occurs. The whole row is omitted, and there would be a glyph name gap for some but not all editions (M and N have gaps, but L has no gap).

With Duplicates				Standardized		
Src L	Src M	Src N		Src L	Src M	Src N
T	T	T	→	T	T	T
1234	1232	1201		1234	1232	1201
U	U	U	→	U	U	U
1235	1233	1202		1235	1233	1202
...	...	...		...	...	...
X	X	X	→	X	X	X
1301	1402	1250		1301	1402	1250
	U	U	→	Y	Y	Y
	1403	1251		1302	1404	1252
Y	Y	Y	→	...	...	...
1302	1404	1252				
...	...	...				

Glyph Name Gap

Figure 3: Row omissions of edition-dependent entries

③ Cell omissions of edition-dependent duplicates

Figure 4 shows the case where the occurrence (the first “V” or second “U”) is found in some or all editions, some but not all instances (the second “U”) are duplicated. Because of non-duplicated instances (the first “V”), the whole row cannot be omitted, but the cell for the duplicated instance (the second “U”) is omitted and leaves a blank at its position. There would be a glyph name gap between the previous and next cell.

With Duplicates				Standardized		
Src L	Src M	Src N		Src L	Src M	Src N
T	T	T	→	T	T	T
1234	1232	1201		1234	1232	1201
U	U	U	→	U	U	U
1235	1233	1202		1235	1233	1202
...	...	...		...	...	...
X	X	X	→	X	X	X
1301	1402	1250		1301	1402	1250
V	U	V	→	V		V
1302	1403	1251		1302		1251
Y	Y	Y	→	Y	Y	Y
1303	1404	1252		1303	1404	1252
...	...	...		...	...	...

Glyph Name Gap

Figure 4: Cell omission of edition-dependent entries

A.2. Relevant Columns of WG2 N5211

To describe the duplicated or unified glyphs, some cropped images of WG2 N5211 are given. For reviewers' convenience, the structure of the chart is briefly described below.

		Real Sung print	Reproductions	Jiguge ed.	font
00001:	孫-魯01上-圖01右-01 陳-魯01上-圖01右-01 N4688v01-01-01 段-魯01上-圖01右-行03-01	干	干	干	干
00002:	孫-魯01上-圖01右-02 陳-魯01上-圖01右-02 N4688v01-01-02 段-魯01上-圖01右-行01-01	干	干	干	干

Figure 5: Structure in WG2 N5211 and relevant columns

The digits under the glyph images are the sequential number of each source (thus, the number in the cell for N4688 differs from TH- glyph name, because TH- glyph names use three different numbering systems; TH- $\backslash d\{5\}$, TH- $X\backslash d\{3\}$, TH- $Y\backslash d\{3\}$.) And the DYC glyph images are taken from kanji-database project (<https://kanji-database.sourceforge.net/dict/swjz/>) including some missing glyphs and unneeded glyphs (glyphs taken from the description text, not heading glyphs), so the number in the DYC cell is not same with the D- $\backslash d\{5\}$ glyph names in Unicode 18.0.0. Anyway, the numbers under the CCZ and QJZ glyphs are consistent with the C- $\backslash d\{5\}$ and K- $\backslash d\{5\}$ glyph names.

A.3. Row Omissions for Universal Duplicates

The items in this section are the items that two quite similar shaped glyphs are placed in all editions of Shuowen Jiezi (THX, CCZ, QJZ and DYC).

A.3.1. U+3D3AB: “吹”

As shown in Figure 6, all editions (THX, CCZ, QJZ and DYC) have two occurrences of “吹”, under the “口” and “欠” radicals.

00921z	孫:港02上葉04右g13 陳:港02上葉08左g03 N4688v02.p11.g07 段:港02上葉16左行06.g01	張	張	張	張	張	張	張	張	張	張	張	張	張	張	張
			00921	00921	00921	00921	00921	00921	00921	01975	11681	09082	00883	00894	00894	00894
06277z	孫:港08下葉04左g05 陳:港08下葉08右g05 N4688v08.p40.g05 段:港08下葉19左行01.g01	張	張	張	張	張	張	張	張	張	張	張	張	張	張	張
			06277	06277	06277	06275	06277	06275	06280	07226	07675	09082	06047	06094	05925	06094

Figure 6: Occurrences of Seal “吹” in SWJZ (WG2 N5211)

In both occurrences, the descriptions are exactly the same; “嘘也。从口从欠。昌垂切。”. Unicode 18.0.0 allocates only one codepoint for these instances, at U+3D3AB, which is the earlier instance under the “口” radical, as shown in Figure 7.

3D3AB

𠵹	22				
𠵹	320	TH-00888	C-00921	K-00894	D-00883
𠵹	5439				

Figure 7: Standardized “吹” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 8.

3E8E9				
𠂔 320	TH-06059	C-06276	K-06069	D-06049
𠂕 6B2F				
3E8EA				
𠂘 320	TH-06061	C-06278	K-06070	D-06051
𠂙 6B28				

Figure 8: Glyph name gap in the range from U+3E8E9 to U+3E8EA



Figure 12: Glyph name gap in the range from U+3D888 to U+3D889

The omitted glyph names under the “又” radical are identified as TH-02078, C-02157, K-02083, and D-02061. Thus, the proposed property can be designed as:

U+3D3DD kSEAL AdditionalSrc TH-02078 C-02157 K-02083 D-02061

A.3.3. U+3D406: “𐑆”

As shown in Figure 13, all editions (THX, CCZ, QJZ and DYC) contain two occurrences of “𠂇”, under the “日” and “𠂇” radicals.

01013e	孫:卷02上葉05右.g32 陳:卷02上葉12左.g10 N4688v02.p18.g08 段:卷02上葉24右行09.g01	呀肝													
			01013	01013	01013	01013	01013	01013	01013	02068		00975	00986	00986	00988
03518:	孫:卷05上葉06左.g04 陳:卷05上葉14右.g07 N4688v05.p16.g14 段:卷05上葉32左行09.g01	群													
			03518	03518	03518	03518	03518	03518	03519	12063	00554	00026	03395	03416	03402 03414

Figure 13: Occurrences of Seal “𠔁” in SWJZ (WG2 N5211)

In both occurrences, the descriptions are semantically equivalent; the earlier entry reads “驚也。从口于聲。況于切。”, and the later entry reads “驚語也。从口从亏。亏亦聲。臣鉉等按、口部已有此重出。于救切。”:

Unicode 18.0.0 allocates only one codepoint for these instances, at U+3D406, which is the occurrence under the “ $\square$ ” radical.



Figure 14: Standardized Seal “𠂇” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 15.



Figure 24: Glyph name gap in the range from U+3D943 to U+3D944

The omitted glyph names are identified as TH-02259, C-02340, K-02266, and D-02244. Thus, the proposed property can be designed as:

U+3DC96 kSEAL_AdditionalSrc TH-02259 C-02340 K-02266 D-02244

A.3.7. U+3DE12: “愷”

As shown in Figure 25, all editions (THX, CCZ, QJZ and DYC) have two occurrences of “愷”, under the “豈” and “心” radicals.

03547	孫:卷05上葉07左.g02 陳:卷05上葉16右.g04 N4688v05.p18.g13 段:卷05上葉37右行08.g01	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷
07601	孫:卷10下葉05左.g13 陳:卷10下葉10左.g10 N4688v10.p44.g01 段:卷10下葉25左行08.g01	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷	愷

Figure 25: Occurrences of Seal “愷” in SWJZ (WG2 N5211)

In both occurrences, the descriptions are similar; the early entry reads “愷：康也。从心、豈，豈亦聲。苦亥切。”，and the later entry reads “愷：樂也。从心豈聲。臣鉉等曰、豈部已有。此重出。苦亥切。”。Unicode 18.0.0 allocates only one codepoint for these instances, at U+3DE12, which is the occurrence under the “豈” radical.



Figure 26: Standardized Seal “愷” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 27.



Figure 27: Glyph name gap in the range from U+3EE39 to U+3EE3A

The omitted glyph names under the “心” radical are identified as TH-07327, C-07601, K-07341, and D-07315. Thus, the proposed property can be designed as:

U+3DE12 kSEAL_AdditionalSrc TH-07327 C-07601 K-07341 D-07315

A.3.8. U+3E923: “𡗗”

As shown in Figure 28, all editions (THX, CCZ, QJZ and DYC) have two occurrences of “𡗗”, under the “欠” and “口” radicals.

01062	孫:卷02上.葉06右.g10 陳:卷02上.葉15右.g04 N4688:v02.p22.g05 段:卷02上.葉28左行02.g01	𡗗	01062	01062	01062	01062	01062	01062	01062	01062	02117	09423	41455	01023	01035	01035	01035
06335	孫:卷08下.葉05右.g33 陳:卷08下.葉11右.g02 N4688:v08.p43.g15 段:卷08下.葉25右.行08.g01	𡗗	06335	06335	06335	06334	06335	06335	06335	06336	07284	15369	41455	06105	06128	05959	06128

Figure 28: Occurrences of Seal “𡗗” in SWJZ (WG2 N5211)

In the earlier occurrence, it is not a headword entry but an alternate form of “𡗗”. The headword entry “𡗗” reads “𡗗：鹿鳴聲也。从口幼聲。”, its accompanying entry for “𡗗” reads “𡗗：𡗗或从欠。”. In the later occurrence, it is a headword entry. The entry reads “𡗗：愁兒。从欠幼聲。”. Unicode 18.0.0 allocates only one codepoint for these instances, at U+3E923, which is the later occurrence under the “欠” radical.



Figure 29: Standardized Seal “𡗗” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 30.



Figure 30: Glyph name gap in the range from U+3D438 to U+3D439

The omitted glyph names under the “口” radical are identified as TH-01029, C-01062, K-01035, and D-01023. Thus, the proposed property can be designed as:

U+3E923 kSEAL_AdditionalSrc TH-01029 C-01062 K-01035 D-01023

A.3.9. U+3EE4E: “𠵹”

As shown in Figure 31, all editions (THX, CCZ, QJZ and DYC) have two occurrences of “𠵹”, under the “口” and “心” radicals.

00932-z	孫:卷02上.葉04右.g24 陳:卷02上.葉09右.g03 N4688:v02.p12.g05 段:卷02上.葉18右.行03.g01	𠵹	00932	00932	00932	00932	00932	00932	00932	01987	11692	12235	00894	00906	00906	00906
07621-z	孫:卷10下.葉05左.g33 陳:卷10下.葉11左.g09 N4688:v10.p45.g06 段:卷10下.葉27左.行05.g01	𠵹	07621	07621	07621	07620	07621	07619	07624	08570	08975	12235	07333	07361	07065	07360

Figure 31: Occurrences of Seal “𠵹” in SWJZ (WG2 N5211)

In the earlier occurrence, it is not a headword entry but an alternate form of “𠵹”. The entry for “𠵹” reads “𠵹：知也。从口折聲。”, its accompanying entry for “𠵹” reads “𠵹：𠵹或从心。”. In the later occurrence, “𠵹” is a headword entry. The entry reads “𠵹：敬也。从心折聲。”.

Unicode 18.0.0 allocates only one codepoint for these instances, at U+3EE4E, which is the later occurrence under the “心” radical.



Figure 32: Standardized Seal “𠵹” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 33.



Figure 33: Glyph name gap in the range from U+3D3B5 to U+3D3B6

The omitted glyph names under the “口” radical are TH-00899 C-00932 K-00906 D-00894.

Thus, the proposed property can be designed as:

U+3EE4E kSEAL_AdditionalSrc TH-00899 C-00932 K-00906 D-00894

A.3.10. U+3EFE3: “湮”

As shown in Figure 34, all editions (THX, CCZ, QJZ and DYC) have two occurrences of “湮”, under the “口” radical and “水” radical.

00914:	孫:卷02上葉04右.g06 陳:卷02上葉08右.g07 N4688:v02.p10.g13 段:卷02上葉16右行05.g01	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮
08016:	孫:卷11上葉04右.g02 陳:卷11上葉09右.g07 N4688:v11.p09.g04 段:卷11上葉57右行02.g01	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮	湮

Figure 34: Occurrences of Seal “湮” in SWJZ (WG2 N5211)

In earlier occurrence, “湮” is not a headword entry but an alternative form of “唾”. The headword entry for “唾” reads “唾：口液也。从口垂聲。湯臥切。”, its accompanying entry for “湮” reads “湮：唾或从水。”. In the later occurrence, “湮” is a headword entry. The entry reads “湮：河津也。在西河西。从水垂聲。土禾切。”.

Unicode 18.0.0 allocates only one codepoint for these instances, at U+3EFE3, which is the later occurrence under the “水” radical.



Figure 35: Standardized Seal “湮” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 36.

It looks like that two glyph names are omitted between U+3D2F7 and U+3D2F8; K-00729 and K-00730. But K-00730 is used at U+3D12E (as shown in Figure 52), so the omitted glyph is K-00729 only in this gap. Thus, the proposed property can be designed as:

U+3D1B6 kSEAL_AdditionalSrc K-00729

A.4.2. U+3D72D: “誤”

As shown in Figure 53, the Sung prints of the DaXu branch and their reproductions (a copy owned by Lu Xinyuan (陸心源舊藏本), a copy owned by Haiyuange library (海源閣舊藏本), and its reproductions like THX, PJG, and CCZ) have duplicated occurrences of “誤”, both under the “言” radical. In both occurrences, the descriptions are exactly same; “謬也。从言吳聲。”.

01809	孫:卷03上.葉05右.g17 陳:卷03上.葉12右.g01 N4688v03.p14.g06 段:卷03上.葉23右.行08.g02	誤																
01809			01809	01809	01809	01809	01809	01809	01809	02787	11769	23013	01728	01752	01758	01752		
01856	孫:卷03上.葉05左.g30 陳:卷03上.葉14右.g10 N4688v03.p17.g05 段:欠	誤																
01856			01856	01856	01856	01856	01856					23013						

Figure 53: Occurrences of Seal “誤” in SWJZ (WG2 N5211)

It should be noted that the duplication of “誤” and “誑” is not same ordered. Their earlier occurrence orders “誤” then “誑”, but the later occurrence orders “誑” then “誤”, as shown in Figure 54.

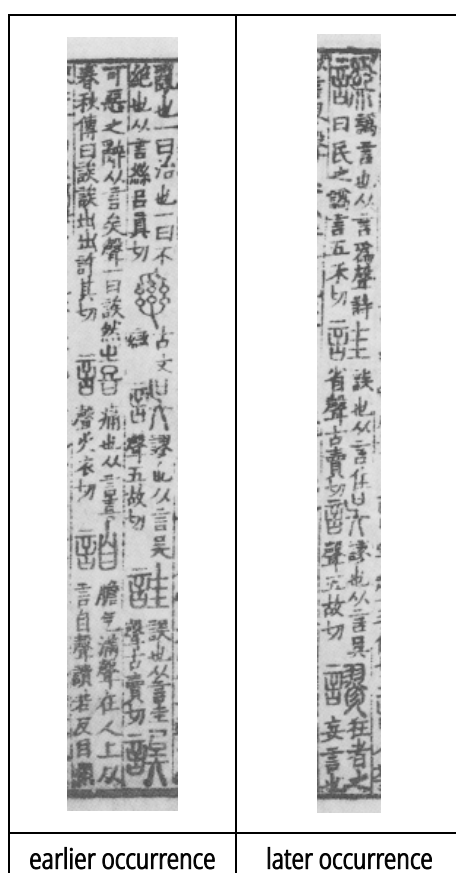


Figure 54: Duplicated occurrences of “誤” and “註” in the Sung print copy owned by the Haiyuange Library (海源閣)

Unicode 18.0.0 allocates only one codepoint for these instances, at U+3D72D, which is the earlier occurrence.



Figure 55: Standardized Seal “誤” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 56.



Figure 56: Glyph name gap in the range from U+3D75A to U+3D75B

A.4.4. U+3E003: “柎”

As shown in Figure 60, the editions in the DaXu branch and DYC edition contain two “𣎵” occurrences. The earlier occurrence is a headword entry under the “木” radical. It reads “𣎵：木也。實如梨。从木尼聲。”. The editions in the XiaoXu branch do not contain this occurrence. The later occurrence is not a headword entry but an alternate form of “屎” under the “木” radical. The headword entry for “屎” reads “屎：糞柄也。从木尸聲。”. And accompanying entry for “𣎵” reads “𣎵：屎或从木。尼聲。臣鉉等曰、𣎵、女氏切、木若梨。此重出。”. The editions in the XiaoXu branch contain this occurrence.

04035:	孫世6上葉02左.g14 陳世6上葉05左.g03 L4688:v06.p06.g13 段世6上葉13右.行04.g01	龍															
			04035	04035	04035	04035	04035	04033	04036	05009	13952	14086	03896				
04311:	孫世6上葉07右.g13 陳世6上葉19右.g09 L4688:v06.p24.g01 段世6上葉52右.行08.g01	龍															
			04311	04311	04311	04311	04311	04309	04312	05284	14083	14086	04172	04193	04179	04193	

Figure 60: Occurrences of Seal “坭” in SWJZ (WG2 N5211)

As shown in Figure 61, Unicode 18.0.0 allocates only one codepoint for these instances, at U+3E003 whose position corresponds to the earlier but non-universal occurrence. It is because this entry is a headword entry. At this codepoint, there is a QJZ source, but this glyph is moved from the later occurrence (see the glyph sequence number of QJZ in Figure 61).

3E003
第 206
梣 67C5

TH-03915 C-04035 K-04193 D-03899

Figure 61: Standardized Seal “梶” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 62.

3E11A
米 206
屎 6758
TH-04190 C-04310 K-04192 D-04174

3E11B
米 206
棚 3B84
TH-04192 C-04312 K-04194 D-04176

Figure 62: Glyph name gap in the range from U+3E11A to U+3E11B

There are gaps in glyph names; TH-04191, C-04311, K-04193, and D-04175 from U+3E11A to U+3E11B. As shown in Figure 61, K-04193 is moved to U+3E003, therefore, the omitted glyphs are identified as TH-04191, C-04311, and D-04175. Thus, the proposed property for U+3E003 can be designed as:

A.4.5. U+3ED9C: “炙”

As shown in Figure 63, the editions in the XiaoXu branch contain two occurrences of “炙”, although the editions of the DaXu branch and the DYC edition contain one occurrence. The earlier occurrence is specific to the XiaoXu editions, which is a headword entry of “炙” under the “火” radical. The later occurrence is an universal headword entry of “炙” as the section head of the “炙” radical.

K07050:	陳-欠	欠																欠	欠	欠
																		07051	06775	07050
07447:	孫-巷10下,葉01右,g08 陳-巷10下,葉01左,g01 N4688v10,s33,g12 段-巷10下,葉02右,行05,g01	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠	欠
			07447	07447	07447	07446	07447	07445	07450	08396	08360	16346	07161	07188	06899	07187				

Figure 63: Occurrences of Seal “炙” in SWJZ (WG2 N5211)

The earlier entry in the XiaoXu editions reads “炙：炙肉也。從肉在火上。臣鍇曰、會意、之射反。又按、炙字別有。部注云、炮肉也、從月在火上、凡炙之屬皆從炙。之石反。此疑誤收。”. The later universal entry reads “炮肉也。從肉在火上。凡炙之屬皆從炙。眞石反。”. Xu Kai suspected the earlier occurrence was mistakenly introduced.

As shown in Figure 64, Unicode 18.0.0 allocates only one codepoint, U+3ED9C, which is the later occurrence under the “炙” radical.

3ED9C				
𤇀 387*				
𤇁 7099	TH-07175	C-07447	K-07187	D-07163
3ED9D				
𤇄 387				
𤇅 2C279	TH-Y149		K-07051	
3ED9E				
𤇆 387				
𤇇 2442C	TH-07176	C-07448	K-07188	

Figure 64: Standardized Seal “灸” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 65.

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 68.

Figure 68: Glyph name gap in the range from U+3EDE3 to U+3EDE4

The omitted glyph name is identified as K-07254. Therefore, the proposed property for U+3EE9D can be designed as:

A.4.7. U+3F2C6: “𪛗”

As shown in Figure 69, some editions in the XiaoXu branch contain two occurrences of “閑”, under the “木” and “門” radicals.¹ The earlier occurrence is a headword entry specific to the XiaoXu branch, it reads “閑：止也。從木門。臣鍇曰、閑猶闌也。《易》曰、「閑有家。」陶潛有《閑情賦》謂「自閑止其情欲也。」候艱反。”。The later occurrence is an universal² headword entry, it reads “閑：闌也。從門中有木。臣鍇按、《易》曰「閑邪存其誠」。陶潛有《閑情賦》謂「闌止其情欲也。」會意。候艱反。”。They are almost identical.

Figure 69: Occurrences of Seal “関” in SWJZ (WG2 N5211)

Unicode 18.0.0 allocates one codepoint for these instances, U+3F2C6, which is the later occurrence under the “門” radical.

¹ In some editions derived from the late print by the Jiguge library in the Qing dynasty has it, while it is regarded as an edition in the DaXu branch. It is because it tried to merge the XiaoXu editions.



Figure 70: Standardized Seal “閑” in the Unicode 18.0.0

The omitted glyph names can be identified by the gaps in the glyph name sequences in Figure 71.



Figure 71: Glyph name gap in the range from U+3E16C to U+3E16D

The omitted glyph is identified as K-04276. Therefore, the proposed property for U+3F2C6 can be designed as:

U+3F2C6 kSEAL_AdditionalSrc K-04276

A.5.2. U+3D43C: “𠬞”

As shown in Figure 75, *DYC* has two “*合*” occurrences, under the “*口*” and “*水*” radicals.

[illegible]

Figure 75: Occurrences of Seal “” in SWJZ (WG2 N5211)

The earlier occurrence is an universal headword entry, it reads “𡗗：山閒陷泥地。从口，从水敗兒。讀若沅州之沅。九州之渥地也，故以沅名焉。”。The later occurrence is not a headword entry but a Guwen form of the entry for “沅”. Other editions uses “沿”; the editions in the DaXu branch read “沿：古文沅。臣鉉等曰、口部已有此重出。”，the editions in the XiaoXu branch read “沿：古文沅如此。臣次立按、說文徐鉉注云口部已有此重出。又本部下文沿字注云、緣水而下。與川反。”。But Duan had “corrected” it by “𡗗”. Duan’s description reads “𡗗：古文沅如此。各本篆作沿。誤。今正。臣鉉等曰。口部已有。此重出。按口部小篆有𡗗。然則鉉時不從水旁也。口部𡗗下曰。山閒陷泥地。從口。從水敗兒。蓋𡗗字在古文則爲沅水、沅州。在小篆則訓山閒陷泥地。如變字在籀文則訓順。在小篆則訓慕。皆同形而古今異義也。古文作𡗗。小篆作沅。𡗗變作究。此同義而古今異形也。”。

Unicode 18.0.0 allocates only one codepoint for these instances, at U+3D43C, which corresponds to the earlier universal occurrence under the “□” radical.

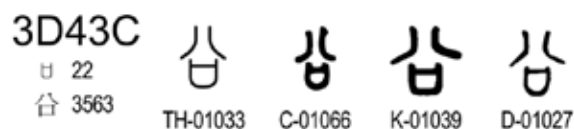


Figure 76: Standardized Seal “” in the Unicode 18.0.0

The omitted glyph name can be identified by the gaps in the glyph name sequences in Figure 77.

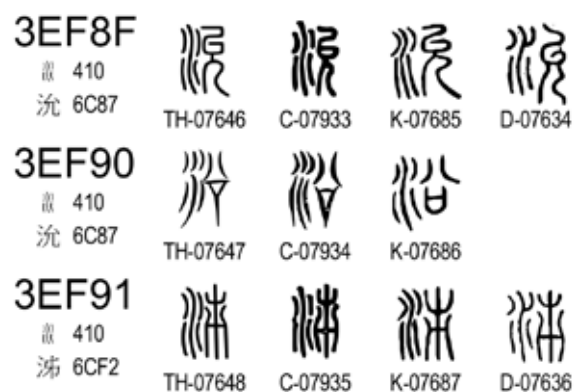


Figure 77: Glyph name gap in the range from U+3EF8F to U+3EF91

3E548 疾 75BE	疾	疾	疾	疾
	TH-05192	C-05363	K-05201	D-05179
3E549 疾 2457A	疾	疾	疾	
	TH-05193	C-05364	K-05203	
3E54A 疾 253F4	疾	疾	疾	疾
	TH-05194	C-05365	K-05202	D-05180
3E54B 痛 75DB	痛	痛	痛	痛
	TH-05195	C-05366	K-05206	D-05182

Figure 83: Glyph name gap in the range from U+3E548 to U+3E54B

The omitted glyph name is identified as D-05181. Thus, the proposed property for U+3D685 can be designed as:

U+3D685 kSEAL_AdditionalSrc D-05181

A.5.5. U+3DE72: “藟”

As shown in Figure 84, the most editions of the DaXu and XiaoXu branches contain two similar pairs of glyphs; “藟”/“藟”, and “藟”/“藟”. Their differences are the contrast of “血” and “皿”. The earlier occurrences under the “艸” radical are alternate forms of “藟”. The headword entry for “藟” reads “藟：酢菜也。从艸沮聲。側魚切。”, its accompanying entry for “藟” reads “藟：或从皿。”, and another accompanying entry for “藟” reads “藟：或从缶。”. The later occurrences under the “血” radical are the headword entry of “藟”, and its alternate form “藟”. The headword entry for “藟” reads “藟：醢也。从血藟聲。側余切。”, its accompanying entry for “藟” reads “藟：或从缶。”. The DYC edition changed “皿” components to “血” components.

00658z	孫:卷01下,葉07左,§12 陳:卷01下,葉20左,§02 N4688:v01.p44,§15 段:卷01下,葉44右,行07,§01(*)	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟
		00658	00658	00658	00658	00658	00658	00658	01714	11551	55232	00637	00639	00639	00639
03641:	孫:卷05上,葉09左,§16 陳:卷05上,葉21左,§02 N4688:v05.p25,§02 段:卷05上,葉51左,行05,§01	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟	藟
		03641	03641	03641	03641	03641	03639	03642	04615	11761	55232	03512	03536	03522	03536

Figure 84: Occurrences of Seal “藟” in SWJZ (WG2 N5211)

Unicode 18.0.0 allocates four codepoints for all instances; “藟”, “藟”, “藟”, and “藟”, as shown in Figures Figure 85 and Figure 86.

3D2A0 𣎵 12 蘊 83F9	TH-00639	C-00657	K-00638	D-00635
3D2A1 𣎵 12 蘊 26F2C	TH-00640	C-00658	K-00639	
3D2A2 𣎵 12 蘊 2709A	TH-00641	C-00659	K-00640	
3D2A3 𣎵 12 荃 8343	TH-00642	C-00660	K-00641	D-00638

Figure 85: Glyph name gap in the range from U+3D2A0 to U+3D2A3

3DE72 𣎵 173 蘊 275CE	TH-03527	C-03641	K-03536	D-03514
3DE73 𣎵 173 蘊 275D8	TH-03528	C-03642	K-03537	D-03515

Figure 86: Standardized Seal “蘊” and “蘊” in the Unicode 18.0.0

Because the DYC edition changed the glyphs for the earlier occurrences, and the resulted glyphs are same with those for the later occurrences, the codepoints for the earlier occurrences have no DYC glyphs. The omitted glyph names are identified as D-00636 for the modified “蘊”, D-00637 for modified “蘊”.

In summary, the proposed property for U+3DE72 can be designed as:

U+3DE72 kSEAL_AdditionalSrc D-00636

A.5.6. U+3DE73: “蘊”

As shown in Figure 87, the DYC edition changed the glyph for “蘊” to “蘊”.

3F9EE				
金 490				
鎔 93DD	TH-10147	C-10531	K-10162	D-10146
3F9EF				
金 490				
櫟 69FE	TH-10148	C-10532		
3F9F0				
金 490				
鎡 947D	TH-10149	C-10533	K-10164	D-10148

The omitted glyph names are identified as K-10163, and D-10147. Thus, the proposed property of U+3E0B3 can be designed as:

A.5.8. U+3F38D: “𠂔”

03928	孫:卷05下葉08右g20 陳:卷05下葉16左g30 N4688v05.p44.g04 段:卷05下葉42左行03.g01	第	          	03928	03926	03928	03928	03926	03926	03929	04902	13240	13100	03792	03818	03804	03816
08935e	孫:卷12上葉06左g01 陳:卷12上葉16右g01 N4688v12.p19.g02 段:卷12上葉37左行05.g01	第	          	08935	08935	08935	08932	08935	08933	08938	12066			08597	08582	08268	08616

The earlier occurrence under the “韋” radical is not a headword entry but an alternative form of “𢶏”. The headword entry for “𢶏” reads “𢶏：收束也。从韋𢶏聲。讀若曾。臣鉉等曰、𢶏側角切。聲不相近未詳。即由切。”, and its accompanying entry for “𢶏” reads “𢶏：𢶏或从秋手。”. The later occurrence under the “手” radical is a headword entry, the entry for “𢶏” reads “𢶏：束也。从手秋聲。《詩》曰：「百祿是𢶏。」即由切。”.

A-36

飲酒義曰。西方者秋。秋之爲言愁也。愁讀爲摯。按韋部韋、收束也。或从要作韋。或从秋手作摯。摯卽摯。然則此篆實爲重出也。” And changed the earlier occurrence to look like a duplicated entry of “摯”.

Unicode 18.0.0 allocates two codepoints for the universal glyph “摯” at U+3F38D (Figure 93) and the DaXu-specific glyph “摯” at U+3DF98 (Figure 94). The codepoint for the DaXu-specific glyph leaves the QJZ and DYC columns as blank.



Figure 93: Standardized Seal “摯” in the Unicode 18.0.0

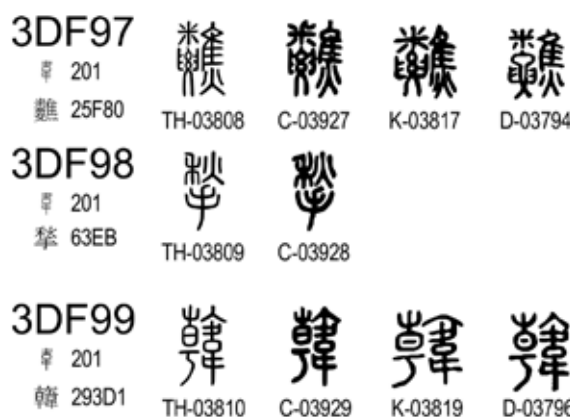


Figure 94: Blanked QJZ and D columns at U+3FF98 for “摯”

The omitted glyph names can be identified as K-03818 and D-03795 from the glyph names at previous and next codepoints. The glyphs K-03818 and D-03795 under the “韋” radical should be displayed by K-08615 or D-8601 at U+3F38D under the “手” radical. Thus, the proposed property can be designed as:

U+3F38D kSEAL_AdditionalSrc K-03818 D-03795

A.5.9. U+3F4EA: “變”

As shown in Figure 95, most editions of SWJZ contain two occurrences “變” under the “女” radical.⁴ The earlier occurrence is a Zhouwen form of “嬌”. The headword entry for “嬌” reads “嬌：順也。从女僇聲。《詩》曰：「婉兮嬌兮。」”，its accompanying entry for “變” reads “變：籀文嬌。”。 The later occurrence is the headword entry, it reads “變：慕也。从女戀聲。”。 Their glyph shapes have

⁴ In the Jiguge edition, two occurrences are unified because its ancestor 五音韻譜 had already unified them.

small structural differences in the DaXu editions. The QJZ edition has similar differences, but the glyphs are transposed⁵. The DYC edition does not have this difference.

09210_	孫:卷12下.葉02左.g15 陳:卷12下.葉06左.g05 N4688v12.p37.g07 段:卷12下.葉14左.行04.g01	變												
09275_	孫:卷12下.葉03左.g11 陳:卷12下.葉10右.g03 N4688v12.p41.g12 段:卷12下.葉21左.行08.g01	變												

Figure 95: Occurrences of Seal “變” in SWJZ (WG2 N5211)

As shown in Figures Figure 96 and Figure 97, Unicode 18.0.0 allocates two codepoints for these instances. The earlier instance is coded at U+3F4A8, and the later instance is coded at U+3F4EA. The omitted glyph name can be identified by the gaps in the glyph name sequences in Figure 96.

3F4A7				
𠂔 443	TH-08872	C-09209	K-08887	D-08864
𠂔 21857				
3F4A8				
𠂔 443	TH-08873	C-09210	K-08952	
變 5B4C				
3F4A9				
𠂔 443	TH-08874	C-09211	K-08890	D-08866
𠂔 59B4				

Figure 96: Glyph name gap in the range from U+3F4A7 to U+3F4A9

3F4EA				
𠂔 443	TH-08938	C-09275	K-08888	D-08930
變 5B4C				

Figure 97: Standardized Seal “變” in the Unicode 18.0.0

As shown in Figure 96, the omitted glyph name for DYC is identified as D-08865. Therefore, the proposed property for U+3F4EA can be designed as:

U+3F4EA kSEAL_AdditionalSrc D-08865

⁵ The SGT edition (述古堂本) and WQS edition (汪啓淑本) in the XiaoXu branch do not have this glyph difference.