

# Universal Character Set

## UCS

ISO/IEC JTC1/SC2/WG2 IRG N2225Draft for review

Date: 2017-07-25<sup>4</sup>






Update from 2017-070-21 Version with red marks.

|               |   |
|---------------|---|
| Source:       | Henry Chan and Qin Lu (IRG Rapporteur)          |
| Title:        | <b>CJK Supplementary Components for IDS Use</b> |
| Distribution: | IRG Members and Ideographic Experts             |
| Medium:       | Electronic                                      |
| Pages:        | 8   |

## 1. Introduction

Based on the IRG Recommendations IRG M48.10, a list of supplementary CJK ideograph components are listed here for use in IRG WS 2017 submissions using IDS with these supplements to increase quality of machine checking of IDS. If this is successful, IRG will consider adding this as a new IWDS series.

The components listed here are mostly extracted from analysis of CJK Ideographs in the URO. The authors consider them to be in common use, relatively speaking. To make identification easy, the components are organized according the first stroke (FS) listed below:

| FS | FS Code  | Name of Stroke                                  | Rep Glyph   | Character Examples         |
|----|----------|---|---|----------------------------|
| 1  | <b>H</b> | Horizontal stroke (“橫”, <i>Héng</i> )           |  | 一 (one), 三 (three)         |
| 2  | <b>S</b> | Vertical stroke (“豎”, <i>Shù</i> )              |  | 卜 (divination), 巾 (thirty) |
| 3  | <b>P</b> | Left slant stroke (“撇”, <i>Piě</i> )            |  | 川 (river), 月 (moon)        |
| 4  | <b>D</b> | Dot stroke (“點”, <i>Diǎn</i> )                  |  | 流 (current), 宝 (precious)  |
| 5  | <b>Z</b> | Turned stroke (“折”, <i>Zhé</i> ) <sup>1</sup> . |  | 彳, 予, 細                    |

Each component is assigned an IDS Code. **The IDS code of the components will start with “&\_” as the component IDS(CIDS) designator sequence**, followed by the FS Code, followed

<sup>1</sup> There are many different forms of the turned strokes and their starting stroke can be any of the above single strokes. No distinction is made in this document.

by the number of strokes of the component, followed by a separator “-”, followed by a serial number in two decimal digits.

苗

For example, the component 苗 will be given the IDS code H7-01. Thus, the CIDS of 曹 is defined as 曹&H7-01 曹.

Each component may have multiple variant glyphs. A head glyph is selected and the other variant glyphs are placed underneath it in each row. The number of strokes is based on the head glyph for the component.

## 2. List of Components

Below is the list of the supplementary components organized according to the first stroke.

### 2.1. FS = 1 (FS Code = H)

| IDS code | Glyph | Description   |
|----------|-------|---|
| H3-01    | 𠃉     | CJK RADICAL PAW FOUR <sup>2</sup><br>Example: bottom of 虐 / 疝<br>= CDP-8DC1         |
| H3-02    | 弋     | CJK COMPONENT C-SIMPLIFIED TRIPLE EARTH<br>Example: 尧<br>= CDP-89D9                 |
| H3-03    | 亻     | CJK COMPONENT CAI<br>Example: outside of 在 / 存<br>= CDP-88F1                        |
| H3-04    | 𠃊     | CJK COMPONENT RAISE ONE<br>Example: bottom of 具 / 巽 / 共                             |
| H5-01    | 巾     | CJK COMPONENT SACK DRAWSTRING WITH STOPPER<br>Example: top of 囊 and 囊<br>= CDP-8DDD |
| H6-01    | 𠃋     | CJK COMPONENT BEAST BODY TWO<br>Example: bottom of 麤焉烏<br>= CDP-8DBF                |
| H6-02    | 𠃌     | CJK COMPONENT LIGATURE DOUBLE ARRIVE<br>Example: top of 晉<br>= CDP-8DAB             |

<sup>2</sup> The name of the components used in this document is only meant as a reference. It has no significance in the IDS sequence. So, its appropriateness will not be discussed for WS 2017. However, if the naming of these components is changed as a result of IRG/WG discussion, update to this document will be done accordingly.

|        |   |   |
|--------|---|---|
| H6-03  | 馬 | CJK COMPONENT PART HORSE<br>Example: outside of 馬<br>= CDP-896A               |
| H7-01  | 𠂔 | CJK COMPONENT LIGATURE DOUBLE EAST<br>Example: top of 曹 / 曹<br>= CDP-8D61     |
| H8-01  | 頁 | CJK COMPONENT LEAF<br>Example: top of 𦉳 / 憂<br>= CDP-8CD4                     |
| H9-01  | 臣 | CJK COMPONENT SIMPLIFIED OVERSIGHT<br>Example: top of 監 / 覽 / 豎<br>= CDP-8CC9 |
| H14-01 | 臨 | CJK COMPONENT OVERSIGHT<br>Example: top of 鑿 / 覽<br>= CDP-8D50                |

## 2.2. FS = 2 (FS Code = S)

| IDS Code | Glyph  | Description   |
|----------|--------|---|
| S2-01    | 冫      | CJK COMPONENT C-SIMPLIFIED MOUND<br>Example: left of 帥 / 歸<br>= CDP-8968                        |
| S4-01    | 収      | CJK COMPONENT C-SIMPLIFIED LEAD<br>Example: top of 堅 / 豎 / 緊 / 腎 / 賢<br>= CDPO-81EF / UTC-02964 |
| S4-02    | 冂      | CJK COMPONENT BOTTOMLESS HOLLOW<br>Example: bottom of 雋<br>= CDP-8DDF                           |
| S5-01    | 𠂔      | CJK COMPONENT C-SIMPLIFIED OVERSIGHT<br>Example: top of 監 / 覽<br>= CDPO-81F0 / UTC-02965        |
| S5-02    | 电      | CJK COMPONENT REPTILE BODY<br>Example: body of 龍, 龜<br>= CDP-8BE8                               |
| S7-01    | 苜      | CJK COMPONENT COCOON<br>Example: outer enclosure of 繭 / 繭                                       |
| S7-02    | 亞<br>曲 | CJK COMPONENT VASE BOTTOM<br>Example: bottom of 壺 / 壺 / 歡                                       |
| S8-01    | 龍      | CJK COMPONENT DRAGON BODY<br>Example: right of 龍  |

|        |             |  |
|--------|-------------|--|
|        |             | = CDP-89B0   |
| S9-01  | 𩺰           | CJK RADICAL J-SIMPLIFIED FROG<br>Example: right of 繩 / 蠅                                 |
| S11-01 | 𩺰<br>𩺰<br>𩺰 | CJK COMPONENT DREAMING PERSON<br>Example: top of 夢 / 薨 / 薈 / 蕘 / 蕘 / 蕘 / 蕘<br>= CDP-8D60 |

### 2.3. FS = 3 (FS Code = P)

| IDS Code | Glyph | Description   |
|----------|-------|---|
| P2-01    | 匚     | CJK COMPONENT LEFT ENCLOSURE<br>Example: left of 兜  |
| P3-01    | 𠃉     | CJK RADICAL PAW THREE<br>Example: left of 印 / 褒 / 褒 / 褒 / 褒 / 處                           |
| P4-01    | 𠃊     | CJK RADICAL V-SIMPLIFIED WIND<br>Example: right of 飢                                      |
| P4-02    | 夕     | CJK RADICAL MEAT TWO<br>Example: top of 炙 / 畜 / 然 / 將                                     |
| P4-03    | 夕     | CJK RADICAL MOON TWO<br>Example: top right of 望<br>= CDP-8B5E                             |
| P4-04    | 𠃋     | CJK COMPONENT LIGATURE DOUBLE HUMAN<br>Example: bottom right of 旅 (J-source) / right of 彳 |
| P4-05    | 𠃌     | CJK COMPONENT SIMPLIFIED PART BIRD<br>Example: outside of 鸟 / 岛                           |
| P5-01    | 𠃍     | CJK COMPONENT LIU<br>Example: top of 留 / top left of 劉<br>= CDP-8C69                      |
| P5-02    | 𠃎     | CJK COMPONENT LOOSE ROCK ONE<br>Example: left of 段<br>= CDP-8CE4                          |
| P6-01    | 𠃏     | CJK COMPONENT BEAST BODY ONE<br>Example: bottom of 鳥<br>= CDP-8C40                        |
| P6-02    | 𠃐     | CJK COMPONENT BEAST HEAD ONE<br>Example: top of 象 / 鳥                                     |
| P6-03    | 𠃑     | CJK COMPONENT COMPLEX PERSON CAVE<br>Example: top of 奧 / 窰                                |

|        |             |  |
|--------|-------------|--|
|        | 𠂇<br>𠂈      | = CDP-8CAF   |
| P7-01  | 𠂉           | CJK COMPONENT FOOD VESSEL TWO<br>Example: left of 卽 / 既                      |
| P7-02  | 𠂊<br>𠂋      | CJK COMPONENT BEAST BODY THREE<br>Example: bottom of 𧢲 / 𧢳 / 𧢴 / 𧢵 / 𧢶 / 𧢷   |
| P7-03  | 𠂌           | CJK COMPONENT PART BIRD<br>Example: outside of 島 / 鳥<br>= CDP-8CBB           |
| P8-01  | 𠂍           | CJK COMPONENT DANCE<br>Example: top of 無 / 舞<br>= CDP-8BBE                   |
| P10-01 | 𠂎           | CJK COMPONENT RAISE TWO<br>Example: enclosure of 輿 / 輿 / 輿 / 輿<br>= CDP-8BA8 |
| P15-01 | 𠂏<br>𠂐<br>𠂑 | CJK COMPONENT COMPLEX POURING VESSEL<br>Example: top of 甕 / 甕                |

#### 2.4. FS = 4 (FS Code = D)

| IDS Code | Glyph | Description   |
|----------|-------|---|
| D4-01    | 𠂒     | CJK COMPONENT OMISSION<br>Example: bottom of 𠂓 / bottom right of 𠂔<br>= CDP-89AE  |
| D5-02    | 𠂕     | CJK COMPONENT C-SIMPLIFIED LEARN<br>Example: top of 学 / 𠂖 / 𠂗 / 𠂘 / 𠂙 / 𠂚 etc.<br>= CDPO-81DD / UTC-02966                                   |
| D6-01    | 𠂛     | CJK COMPONENT C-SIMPLIFIED CHAOS<br>Example: top of 恋 / 𠂜 / 𠂝.<br>Note: use 4EA6 亦 instead where appropriate, such as 5955 𠂞<br>= UTC-02967 |
| D6-02    | 𠂟     | CJK COMPONENT VIRTUOUS<br>Example: left of 朗 / 朗, left hand side form of 良<br>= CDP-8C42  |

|        |   |  |
|--------|---|--|
| D7-01  | 宀 | CJK COMPONENT ARCHITECTURAL STRUCTURE<br>Example: top of 亮 / 亭<br>= CDP-8C4D |
| D7-02  | 宀 | CJK COMPONENT GUEST<br>Example: top of 寶<br>= CDP-8D74                       |
| D7-03  | 廌 | CJK COMPONENT BEAST HEAD TWO<br>Example: outside of 鹿 / 廌<br>= CDP-8D56      |
| D10-01 | 廸 | CJK COMPONENT EVEN<br>Example: top of 廸 / 齊<br>= CDP-8DEB                    |

### 2.5. FS = 5 (FS Code = Z)

| IDS Code           | Glyph | Description   |
|--------------------|-------|---|
| Z2-01              | 勹     | CJK COMPONENT RIGHT ENCLOSURE<br>Example: right of 兜 / top right of 段   |
| Z2-02              | 勹     | CJK COMPONENT LEFT HAND ONE<br>Example: left of 収   |
| Z2-03              | 匕     | CJK COMPONENT LEFT HAND TWO<br>Example: right of 𠂇<br>= CDP-88B5  |
| Z2-04              | 𠂇     | CJK COMPONENT SIMPLIFIED PART HORSE<br>Example: outside of 马  |
| Z2-05              | 𠂇     | CJK COMPONENT RIGHT HAND TWO<br>Example: right top of 祭 / 登<br>= CDP-8C4E   |
| Z3-01              | 冫     | CJK COMPONENT RIGHT HAND ONE<br>Example: top of 冫 / top right of 侵 / bottom of 𠂇<br>= CDP-8BAB                                      |
| Z3-02              | 廴     | CJK COMPONENT HOLD ONE<br>Example: right top of 恐 / 鞏 etc   |
| Z3-03 <sup>3</sup> | 丸     | CJK COMPONENT HOLD TWO (variant of 廴)<br>Example: right of 執 / 執 / 狻 / 秣<br>Note: use 4E38 丸 (wan2) where appropriate, e.g. 7D08 紉. |
| Z4-01              | 𠂇     | CJK COMPONENT TILTED HEAD<br>Example: bottom of 吳<br>= CDP-8CD5   |

<sup>3</sup> Refer to Appendix A: Rationale for Inclusion of Z3-03.

|       |   |  |
|-------|---|--|
| Z4-02 | 𠄎 | CJK COMPONENT RAN<br>Example: left of 那<br>= CDP-89FA                    |
| Z5-01 | 艮 | CJK COMPONENT FOOD VESSEL ONE<br>Example: left of 即 / 既<br>= CDP-8B7C    |
| Z5-02 | 𠄎 | CJK COMPONENT LOOSE ROCK TWO<br>Example: left of 段<br>= CDP-8C7A         |
| Z7-01 | 鼠 | CJK COMPONENT BEAST BODY THREE<br>Example: bottom of 鼠 / 鼯<br>= CDP-8D46 |

### 3. Additional Components for Reference

The table below also supplies some coded ideographs which can serve as common component variants. For easy IDS checking, the base character as a component should be used in the IDS sequence to represent the variant glyph directly.

| Base Character | Variant Glyph | Descriptions   |
|----------------|---------------|--|
| U+2B94E        | 𠄎             | Example: right of 𠄎 / 𠄎                                    |
| U+53C0<br>𠄎    | 𠄎             | Example: top of 𠄎 / 𠄎<br>= CDP-8BD0                        |
| U+5C6F<br>屯    | 𠄎             | Example: right of 𠄎 / 𠄎 / 𠄎<br>= CDP-8AF0                  |
| U+6B62<br>止    | 𠄎             | Example: bottom of 𠄎 / 𠄎<br>= CDP-8962                     |
| U+9FB0<br>𠄎    | 𠄎             | Example: bottom of 𠄎 / 𠄎 / 𠄎<br>= CDP-8961                 |
| U+3B30<br>𠄎    | 𠄎             | Example: right of 𠄎 / 𠄎 / 𠄎<br>= CDP-89E7                  |
| U+4E23<br>𠄎    | 𠄎<br>𠄎        | Example: top right of 𠄎, right of 𠄎 / 𠄎 / 𠄎<br>= CDPO-85FD |





## Appendix A: Rationale for Inclusion of Z3-03

Z3-03 丸 is included as a separate component from U+4E38 丸. First, they are considered different components in the Taiwan and Hong Kong locale. For characters whose etymology is U+4E2E 𠄎 (to hold), the glyph shape Z3-03 with the dot protruding is used. For characters whose etymology is U+4E38 丸 (small round object), the T/H glyph shape of U+4E38 with the dot not protruding is used.

This is a similar case to the existing dis-unified radicals 2EC0 卄 CJK RADICAL GRASS THREE from 2EBE 卅 CJK RADICAL GRASS ONE, and 2EBC 月 CJK RADICAL MEAT from 2F49 月 KANGXI RADICAL MOON. The difference in the shape is reflective of a different etymology and hence “non-cognateness” of characters in the Taiwan and the Hong Kong locales.

Second, these normally unifiable differences in components could also reflect a difference in meaning in certain context for other locales, and not just as a simple typographical variation. For example, PRC selectively applies such typographical variations as a standardized difference for a range of non-cognate characters. The difference between U+2EC0(卄) and U+2EBE(卅) is used to distinguish between the two characters U+830D 苟 and U+82DF 苟 in the PRC locale.

Thus, the inclusion of Z3-03 component is also meaningful and useful for IDS composition.

1. On general principles

According to the document, the component list is given with the intention “to increase quality of machine checking of IDS”, along with that “improve the quality of IDS database” (IRGN2220) and “enhance the ability to use intuitive Ideographic Description Sequences” (IRGN2204). However, we find the current status of the list not optimally designed in light of its ideal and motivation. As we observed, the issues in general fall under several types:

(Actual review for individual components follows in the next section.)

a) Etymological separation of similar shapes

Although it is crucial for encoding ideographs to refer to their historical background, emphasizing etymological correctness in IDS decomposition may sometimes complicate comparison and matching workflow, as graphically intuitive division of a modern ideograph is not necessarily backed by its etymology. For example, 香 can be easily described like 𠂇 禾日, where the bottom part does not derive from 日 nor 曰 (but 甘). Therefore we should be restrictive from loading etymological semantics on components without radical use.

b) Unification criteria

We find that the criteria which divide some components in the main list (§2) and those in reference components list to be unified (§3) is not enough clear. Some §3 components do have minimal pairs in current standard while some of §2 do not. Any justification or instruction will be appreciated. Supposedly unifiable components are mentioned in the next section.

c) Insufficient isolation of compound shapes

Some shapes are seemingly too big to be monolithic components. While they may have been understandably induced from the set in BMP, further decomposition would contribute to flexibility, future utility, and fuller coverage, with which the original objective of describing commonest characters remains unaffected.

2. Individual issues

|        |   |   |
|--------|---|---|
| H9-01  | 𠂇 | CJK COMPONENT SIMPLIFIED OVERSIGHT<br>Example: top of 監 / 覽 / 𠂇<br>= CDP-8CC9 |
| H14-01 | 𠂇 | CJK COMPONENT OVERSIGHT<br>Example: top of 鑿 / 𠂇<br>= CDP-8D50                |

These shapes are easily decomposable (1.c). Singling out the right half of H9-01 will allow more usability and coverage. H14-01 seems redundant.

|       |   |   |
|-------|---|---|
| S4-01 | 収 | CJK COMPONENT C-SIMPLIFIED LEAD<br>Example: top of 坚 / 竖 / 紧 / 肾 / 贤<br>= CDPO-81EF / UTC-02964 |
|-------|---|---|

This shape is easily decomposable (1.c). The left part can be described using S2-01, which is also a simplified shape that etymological distinction are not pointful (1.a).

|       |   |  |
|-------|---|--|
| S5-01 | 𠂆 | CJK COMPONENT C-SIMPLIFIED OVERSIGHT<br>Example: top of 监 / 宽<br>= CDPO-81F0 / UTC-02965 |
|-------|---|--|

This shape is easily decomposable (1.c). It shares left part with S4-01 (S2-01) and right part with H9-01.

|       |   |   |
|-------|---|---|
| P4-02 | 夕 | CJK RADICAL MEAT TWO<br>Example: top of 炙 / 畜 / 然 / 將         |
| P4-03 | 𠂇 | CJK RADICAL MOON TWO<br>Example: top right of 望<br>= CDP-8B5E |

The distinction between the two shapes may be blurred when they do not serve for meaning (1.a). They are however still useful considering the universal recognition of MEAT-MOON distinction among ideographic regions.

|       |   |  |
|-------|---|--|
| P5-01 | 𠂈 | CJK COMPONENT LIU<br>Example: top of 留 / top left of 劉<br>= CDP-8C69 |
|-------|---|--|

This shape is decomposable (1.c). Singling out the left half will allow more usability and coverage.


|       |   |   |
|-------|---|---|
| P7-01 | 自 | CJK COMPONENT FOOD VESSEL TWO<br>Example: left of 卽 / 既 |
|-------|---|---|

|       |   |   |
|-------|---|---|
| Z5-01 | 𠂉 | CJK COMPONENT FOOD VESSEL ONE<br>Example: left of 卽 / 既<br>= CDP-8B7C |
|-------|---|---|


Needs more clarification about distinction and/or use cases regarding the two shapes (1.b).

|        |             |   |
|--------|-------------|---|
| P15-01 | 𠂊<br>𠂋<br>𠂌 | CJK COMPONENT COMPLEX POURING VESSEL<br>Example: top of 甕 / 甕 |
|--------|-------------|---|



This shape is decomposable (1.c). Making a component without 同 in the middle (=CDP-8B64) will enable to include other shapes in place of it.

|       |   |   |
|-------|---|---|
| D6-01 |  | CJK COMPONENT C-SIMPLIFIED CHAOS<br>Example: top of 恋 / 婁 / 峦.<br>Note: use 4EA6 亦 instead where appropriate, such as 5955 奕<br>= UTC-02967 |
|-------|---|---|


Distinction with 亦 is not meaningful being a simplified shape (1.a, 1.b).

|       |   |  |
|-------|---|--|
| D7-02 |  | CJK COMPONENT GUEST<br>Example: top of 寶<br>= CDP-8D74 |
|-------|---|--|

This shape is easily decomposable (1.c). Singling out the bottom half will allow more usability and coverage.



|                    |   |   |
|--------------------|---|---|
| Z3-02              |  | CJK COMPONENT HOLD ONE<br>Example: right top of 恐 / 鞏 etc   |
| Z3-03 <sup>3</sup> |  | CJK COMPONENT HOLD TWO (variant of 𠂇)<br>Example: right of 執 / 執 / 狽 / 秣<br>Note: use 4E38 丸 (wan2) where appropriate, e.g. 7D08 絜. |

The distinction between these shapes and existing characters such as 𠂇, 丸, 凡, 凡 etc. may be blurred when they do not serve for meaning (1.a). Addition of these components would not likely to provide much more usefulness.

|       |   |   |
|-------|---|---|
| Z4-01 |  | CJK COMPONENT TILTED HEAD<br>Example: bottom of 吳<br>= CDP-8CD5 |
|-------|---|---|

This glyph is not significantly different from 𠂉, especially compared with unification examples, thus should be subject to unification (1.b).

### 3. Proposal for addition

| Shape   | Description   | Examples        |
|---|---|-----------------|
|  | A shape mainly derives from 𠂊 (U+233B5).<br>= CDP-8D52                  | 索, 孛, 寔 etc.    |
|  | A shape mainly derives from the upper part of 𠂋 (U+41C2).<br>= CDP-8BAE | 帝, 商, 商, 产 etc. |

### 4. On notation

The suggested format of provisional components in IDS is such as 𠂊H7-01 𠂊. We however

propose more formally (and visually) significant notation that using HTML entity reference format, such as `&H7-01;`. It also agrees with the convention in some of Taichi Kawabata's documents.

The advantages are:

- more visually segmented (e.g. `&H7-01;&H3-04;`)
- easier to define in program
- easier to detect mistypes with program

The disadvantages are:

- type more letters
- possibly more mistypes

## Comments on N2225 (CJK Supplementary Components for IDS Use), plus Suggestions for Additional List of Components and New Additional Components

Author: [Leroy Vargas](#)

Type: Individual Contribution to IRG #49

Date: 2017/08/27

References: [IRG N1183](#); [IRG N2117](#); [IRG N2204](#); [IRG N2225](#) Feedback ([Additional Components from Lee Collins](#); [SAT Comments](#)); [IRG N2242](#); [IRG N2243](#); [Unicode Standard Annex #45](#)




---




### Comments on SAT's individual issues:

On 2017/07/24, SAT raised the issue of components that can be IDS-decomposed into other sub-components (general principle 1.c in their comments), of which some or all are already encoded as CJK ideographs. This may be true, but only in the context of deep, maximal IDS decomposition, which results in a sometimes very lengthy IDS sequence that may be easy for machine checking, but hard for eye-ball checking or for checking whether two similar sequences are cognate/non-cognate and/or share/differ in phonetic value.



Principle 2.1 of IRG N1183 (Minimal division) recommends not to “divide too much”, stressing that further division can be easily achieved through machine recursion. The principal idea behind N2204 by Henry CHAN and N2225 by Chan and Prof. Qin LU is to improve readability and quality of IDS sequences – which may imply minimizing the lengths of IDS sequences (by employing as few IDCs as possible). Also, further decomposing a decomposable component may do away with the overall component's phonetic value –and thus, the intended meaning–, thus making it harder to check whether two similar components with slightly-different subcomponents but potentially equal phonetic value might share the same meaning and thus be unified or not. So, while it is technically desirable to single out a subcomponent of a complex component by IDS decomposition so that it can be reused by other components, in practice it is also desirable to keep the overall component with a fixed phonetic and/or meaning so as to optimize human-readable IDS generation, especially when two or more components sharing the same subcomponent are semantically related to each other in the context of variance or make up a traditional/simplified pair (or, in some instances, a traditional/J-simplified/C-simplified trio, or a traditional/C-simplified/V-simplified trio, etc.).




Hereby, the author of this feedback document suggests instead splitting N2225's list of unencoded supplementary components (currently section “2. List of Components”, aliased §2 from here on) into two sections: one listing non-IDS-decomposable components (called atomic components) and another one listing IDS-decomposable components (called non-atomic components) that may or may not contain any of the atomic subcomponents from the first section. §2 can be renamed to “2. List of **Atomic Components**” while current section “3. Additional Components for Reference” (§3o) can be renumbered to “**4. Additional Components for Reference**” (§4); the new list of decomposable components can become new section “**3. List of Non-Atomic Components**” (§3n to distinguish it from §3o which is itself renamed to §4), as can be previewed at the near end of this document.

As for the individual issues: consider SAT’s first one concerning components H9-01, H14-01, & S5-01 shown below. SAT argued that rather than proposing these three separate components, it was enough to merely single out the **right-hand subcomponent** of both H9-01 & S5-01, which is also the **top-right subcomponent** of H14-01, as a proposed component (to be given IDS code **P3-04** by this author) that could be reused by other components via IDS composition. However, following the same decomposition logic suggested by SAT, it appears that the same subcomponent **P3-04**, circled in **red** below, can be itself decomposed into  (or its PRC/Vietnamese variant , as used in S5-01). (Its **traditional form** found in H14-01 and circled **green** below, which is given here IDS code **P8-03**, also contains **P3-04** and decomposes into )

| IDS Code | Glyph   | Description  |
|----------|---|--|
| H9-01    |  | CJK COMPONENT SIMPLIFIED OVERSIGHT<br>Example: top of 監 / 覽 / 豎<br>= CDP-8CC9            |
| H14-01   |  | CJK COMPONENT OVERSIGHT<br>Example: top of 鑿 / 覽<br>= CDP-8D50                           |
| S5-01    |  | CJK COMPONENT C-SIMPLIFIED OVERSIGHT<br>Example: top of 監 / 覽<br>= CDPO-81F0 / UTC-02965 |

However, further decomposing **P3-04** would hide the fact that it can assume at least two variant shapes, each differing by the type of stroke at the bottom (‘一’ versus ‘丿’, both which are non-cognates and thus cannot be individually unified), so it is best to keep **P3-04** as an “atomic component” in spite of its obvious decomposition(s). Also, it is in the best interest to keep **P8-03** as an atomic component in spite of the presence of **P3-04** in it, because both **P8-03** & **P3-04** appear to form a traditional/simplified pair, just like 龠 & 𪛗; they can both be allocated into §2 like this:


| IDS Code     | Glyph   | Description  |
|--------------|---|--|
| <b>P3-04</b> |  | <b>CJK COMPONENT SIMPLIFIED PART OVERSIGHT</b><br>Example: top right corner of 監 / 覽<br>Component proposed by SAT<br>Variant form is for PRC, also requested for Vietnamese by Lee COLLINS |
| <b>P8-03</b> |  | <b>CJK COMPONENT PART OVERSIGHT</b><br>Example: top right of 鑿 / 覽<br>New component proposed by this author  |


Extending the same reasoning to the original components H14-01, H9-01, & S5-01, which all share a common meaning (“oversight”) and a respective traditional/J-simplified/C-simplified form relationship, the author of this feedback believes that rather than just deleting these three components for the mere reason that each contains **P3-04**, they can simply be relocated to §3n, with their respective minimal IDS decompositions , , & ; recorded alongside for machine-recursive decomposition purposes.

In another issue, SAT also pointed out the decomposability of S4-01, which happens to make use of S2-01, already present in §2 and also used by S5-01 above. Following the same logic laid out above, this author also suggests moving S4-01 to §3n (with decomposition  $\square \&S2-01; \text{又}$ ).




| IDS Code | Glyph   | Description   |
|----------|---|---|
| S4-01    |  | CJK COMPONENT C-SIMPLIFIED LEAD<br>Example: top of 坚 / 竖 / 紧 / 肾 / 贤<br>= CDPO-81EF / UTC-02964 |


SAT also pointed out the decomposability of P5-01, singling out its **left-half subcomponent** as a candidate proposal (this author suggests IDS code P3-05, originally treated by N1183's appendix text file [IRGN1183A\\_ids\\_irg.txt](#) as a variant of U+2008E 𠄎), but did not provide evidence of other ideographs using P3-05 to support such singling-out. Nevertheless, this author concurs with SAT in adding P3-05 to §2 and relocating P5-01 to §3n (with decomposition  $\square \&P3-05; \text{刀}$  for recursion purposes).

| IDS Code | Glyph   | Description  |
|----------|---|--|
| P5-01    |  | CJK COMPONENT LIU<br>Example: top of 留 / top left of 劉<br>= CDP-8C69 |

| IDS Code | Glyph   | Description   |
|----------|---|---|
| P3-05    |  | CJK COMPONENT PART LIU<br>Example: top left of 留<br>Component proposed by SAT |

Concerning P15-01, SAT suggested “making a component without 同 in the middle” (to be given here IDS code P9-01), but again did not provide evidence of ideographs using P9-01 but not 同. So, although it may be a good idea to propose P9-01 at §2 (in the same spirit as the similarly-looking P10-01), this author still suggests moving P15-01 to §3n (with decomposition  $\square \&P9-01; \text{同}$ ) rather than dismissing it.

| IDS Code | Glyph   | Description   |
|----------|---|---|
| P15-01   | <br><br> | CJK COMPONENT COMPLEX POURING VESSEL<br>Example: top of 甬 / 甬 |

| IDS Code | Glyph   | Description   |
|----------|---|---|
| P9-01    |  | CJK COMPONENT PART POURING VESSEL<br>Example: top enclosure of 甬 / 甬<br>New component proposed by SAT |



### Comments on SAT's proposal for addition:

The two components proposed for addition by SAT, using the same logic used by SAT to question many of the decomposable components like those mentioned above, can also be decomposed into already-encoded subcomponents, but since they do not exist as standalone encoded ideographs and they seem to have an implied semantic and/or phonetic value, they can too be listed in §3n with IDS codes, decompositions, and suggested provisional names given below:

#### 3.1. FS = 1 (FS Code = H)

| IDS code | Glyph | Decomposition | Description   |
|----------|-------|---------------|---|
| H4-02    | 𠂇     | 𠂇             | CJK COMPONENT COMET TOP<br>Example: top of 索, 字, 寔 etc.<br>= CDP-8D52 |

#### 3.4. FS = 4 (FS Code = D)

| IDS code | Glyph | Decomposition | Description   |
|----------|-------|---------------|---|
| D4-03    | 𠂉     | 𠂉             | CJK COMPONENT ANCIENT FAULT TOP<br>Example: top of 辛, 帝, 商, 商, 产 etc.<br>= CDP-8BAE |

### Comments on additional components proposed by Lee Collins:

On 2017/07/21, Lee COLLINS suggested to add 17 new components to the list originally formulated in N2225 §2. Following the same logic used by SAT above, 11+ of these components can be decomposed into smaller subcomponents (of which one is already proposed as standalone ideograph in IRG Working Set 2015 and another one is proposed new by this author [although it might be unifiable with a similar, already-encoded ideograph]); also, one component is already proposed standalone in WS 2015, and another one is a variant of an already-encoded ideograph:

#### 2.1. FS = 1 (FS Code = H)

| IDS code | Glyph | Description + comments, suggestions, & additions by Vargas   |
|----------|-------|--|
| H9-02    | 𠂊     | CJK COMPONENT 78CA VARIANT ONE<br>WS2015 sequence number 02833 (UTC-01039), can be decomposed into 𠂊石双, including this as a component into N2225 list(s) might seem questionable or at best debatable (Does this belong to §3n or §4?) |
| H9-03    | 𠂋     | CJK COMPONENT 8944 VARIANT ONE<br>Example: bottom right of 爆<br>Can be decomposed into 𠂋𠂉衣, suggest moving to §3n and renaming it CJK COMPONENT HELP VARIANT ONE.  |
| H11-01   | 𠂌     | CJK COMPONENT 6F22 RIGHT HALF<br>Example: right of 漢, unifiable variant of U+26C29 莫<br>Suggest moving this to §4 as an "additional component for reference".  |
| H14-02   | 𠂍     | CJK COMPONENT 21570 VARIANT ONE<br>Example: right of 𠂍<br>Can be decomposed into 𠂍𠂉𠂉𠂉米八文, suggest moving to §3n and renaming it CJK COMPONENT ZONG VARIANT ONE.  |

**Comment [A1]:** Changed font to Arial for consistency with all others

**Comment [A2]:** Also, Collins did not provide evidence of ideographs using this as a component; only the standalone ideograph UTC-01039 has been attested by this author.

2.2. FS = 2 (FS Code = S)

| IDS code | Glyph | Description + comments, suggestions, & additions by Vargas  |
|----------|-------|---|
| S4-03    | 𠂇     | CJK COMPONENT 2A7FD VARIANT<br>Example: right of 𠂇<br>Can be decomposed into 𠂇𠂇, suggest moving to §3n.   |
| S8-02    | 𠂈     | CJK COMPONENT 7F85 VARIANT ONE<br>Example: left of 𠂈, right of 𠂈<br>Can be decomposed into 𠂈𠂈, suggest moving to §3n and renaming it<br>CJK COMPONENT GAUZE VARIANT ONE.  |
| S8-03    | 𠂉     | CJK COMPONENT 7F85 VARIANT TWO<br>Example: right of 𠂉, bottom of 𠂉<br>Can be decomposed into 𠂉𠂉, suggest moving to §3n and renaming it<br>CJK COMPONENT GAUZE VARIANT TWO.  |
| S8-04    | 𠂊     | CJK COMPONENT RAISE THREE<br>Example: right of 𠂊<br>Enclosure of S11-02; similar to S7-03; maybe unifiable w/U+2D048 𠂊?   |
| S11-02   | 𠂋     | CJK COMPONENT 8209 VARIANT<br>Example: right of 𠂋<br>Can be decomposed into two unencoded components: 𠂋⑧③, where ⑧ is the component proposed above as S8-04 (unifiable with U+2D048 𠂊?) and ③ is WS2015 sequence num. 00002 (UTC-01671), which is similar to U+20006 𠂋; suggest moving this component to §3n and renaming it to<br>CJK COMPONENT RAISE VARIANT. |

**Comment [A3]:** If indeed unifiable, please move this proposed component to §4 as an "additional component for reference" with base character U+2D048 𠂊.

2.3. FS = 3 (FS Code = P)

| IDS code | Glyph | Description + comments, suggestions, & additions by Vargas  |
|----------|-------|---|
| P4-06    | 𠂌     | CJK COMPONENT 98ED RIGHT HALF<br>Example: right of 𠂌<br>Can be decomposed into 𠂌𠂌, suggest moving to §3n and renaming it<br>CJK COMPONENT ORDER RIGHT HALF. |
| P8-02    | 𠂍     | CJK COMPONENT 8200 VARIANT<br>Example: right of 𠂍<br>Can be decomposed into 𠂍𠂍, suggest moving to §3n and renaming it<br>CJK COMPONENT LADLE VARIANT.       |

2.4. FS = 4 (FS Code = D)

| IDS code | Glyph | Description + comments, suggestions, & additions by Vargas  |
|----------|-------|---|
| D8-01    | 𠂇     | CJK COMPONENT 7FA9 VARIANT ONE<br>E.g: right of 𠂇, 𠂇, 𠂇<br>Can be decomposed into 𠂇𠂇𠂇, suggest moving to §3n and renaming it CJK COMPONENT RIGHTEOUS VARIANT ONE.   |
| D9-01    | 𠂈     | CJK COMPONENT TRIGRAM FIVE VARIANT ONE<br>E.g: left of 𠂈<br>Can be decomposed into 𠂈𠂈𠂈, suggest moving to §3n.  |
| D13-01   | 𠂉     | CJK COMPONENT 8208 VARIANT ONE<br>Example: left of 𠂉, bottom of 𠂉<br>Can be decomposed into 𠂉𠂉𠂉 & D4-01; 同 & H3-04; where D4-01 is being repurposed as an enclosing component (see author's comment on such component two pages below); suggest moving to §3n & renaming it to CJK COMPONENT PROSPER VARIANT ONE. |

Comment [A4]: For consistency with all other additions, change this to "Example"

Comment [A5]: Source of this ideograph? I'd appreciate it if sources were included here for Collins' three unencoded example ideographs.

Comment [A6]: Ditto as in LV4 comment above.

Also, this author identified three non-decomposable components by Collins whose provisional hex-numeric names could be renamed to more meaningful ones:

2.3. FS = 3 (FS Code = P)

| IDS Code | Glyph | Description + comments, suggestions, & additions by Vargas   |
|----------|-------|--|
| P3-03    | 𠂊     | CJK COMPONENT 4E2A VARIANT ONE<br>Example: right of 𠂊<br>Derived from 𠂊, not 4E87<br>Suggest renaming it CJK COMPONENT SINGLE VARIANT ONE. |

2.4. FS = 4 (FS Code = D)

| IDS Code | Glyph | Description + comments, suggestions, & additions by Vargas   |
|----------|-------|--|
| D4-02    | 𠂋     | CJK COMPONENT 5198 VARIANT ONE<br>E.g: bottom right of 𠂋<br>Suggest renaming it CJK COMPONENT MOVE ON VARIANT ONE. |
| D5-03    | 𠂌     | CJK COMPONENT 5198 VARIANT TWO<br>E.g: right of 𠂌<br>Suggest renaming it CJK COMPONENT MOVE ON VARIANT TWO.        |

Comment [A7]: Source of this ideograph?

Comment [A8]: Ditto?

**Comments on some original components + author's own additions:**

The author of this feedback document is also adding **his own comments and suggestions** to some decomposable components that may have been overlooked by SAT and other involved parties, along with proposing **some new additions of his own** after analyzing some of the decomposables in below and comparing them to other already-encoded ideographs that contained the same component(s):



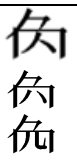
**2.1. FS = 1 (FS Code = H)**

| IDS code | Glyph | Description + comments, suggestions, & additions by Vargas   |
|----------|-------|--|
| H2-01    | 与     | CJK COMPONENT SIMPLIFIED AND<br>Example: enclosure of 与, middle enclosure of 烏<br>Similar in shape to U+20001 与, but not cognate; enclosure of H6-01.  |
| H3-04    | 六     | CJK COMPONENT RAISE ONE<br>Example: bottom of 具 / 巽 / 共<br>Can be decomposed into 日一八, suggest moving to §3n.  |
| H6-01    | 𠂇     | CJK COMPONENT BEAST BODY TWO<br>Example: bottom of 烏 / 焉 / 烏<br>= CDP-8DBF<br>Can be decomposed into 𠂇②..., where ② is a new component to be proposed as H2-01; suggest creating new component H2-01 above at §2 and moving this overall component to §3n. |
| H6-02    | 𠂇     | CJK COMPONENT LIGATURE DOUBLE ARRIVE<br>Example: top of 晉<br>= CDP-8DAB<br>Can be decomposed into 日一从一, suggest moving to §3n.   |








**2.2. FS = 2 (FS Code = S)**

| IDS Code | Glyph       | Description + comments, suggestions, & additions by Vargas  |
|----------|-------------|---|
| S3-01    | 𠂇           | CJK COMPONENT ANCIENT PAIR<br>Example: enclosure of 𠂇, bottom enclosure of 𠂇<br>Bottom part of S7-01.   |
| S7-01    | 𠂇           | CJK COMPONENT COCOON<br>Example: outer enclosure of 𠂇 / 𠂇<br>Can be decomposed into either 日𠂇③ or 日𠂇③, where ③ is a new component to be proposed as S3-01; suggest creating new component S3-01 above at §2 and moving this overall component to §3n. |
| S11-01   | 𠂇<br>𠂇<br>𠂇 | CJK COMPONENT DREAMING PERSON<br>Example: top of 夢 / 薨 / 薨 / 薨 / 薨 / 薨<br>= CDP-8D60<br>Can be decomposed into either 日𠂇𠂇 or 日𠂇𠂇 (depending on desired variant), suggest moving to §3n.   |

**2.3. FS = 3 (FS Code = P)**

| IDS Code | Glyph   | Description + comments, suggestions, & additions by Vargas   |
|----------|---|--|
| P2-02    |  | <b>CJK COMPONENT OMISSION RIGHT HALF</b><br>Example: top right of 樂 / 蕤 (D13-01)<br>Right-half subcomponent of D4-01, especially when used as enclosure.   |
| P6-01    |  | <b>CJK COMPONENT BEAST BODY ONE</b><br>Example: bottom of 𪛗<br>= CDP-8C40<br>Can be decomposed into 𠃉 勹, suggest moving to §3n.  |
| P6-03    |  | <b>CJK COMPONENT COMPLEX PERSON CAVE</b><br>Example: top of 𠃉 / 𠃊<br>= CDP-8CAF<br>Basic form can be decomposed into 𠃉 𠃊; variant forms must check whether bottom-most component shows same variation in isolation. Anyway, suggest moving to §3n. |

**2.4. FS = 4 (FS Code = D)**


| IDS Code | Glyph  | Description + comments, suggestions, & additions by Vargas   |
|----------|--|--|
| D4-01    | <br><br> | <b>CJK COMPONENT OMISSION</b><br>Example: bottom of 裏 / bottom right 涉 / enclosure of 樂 / 蕤 / 尾(尾)<br>= CDP-89AE<br>Can be decomposed into either 𠃉 ② (where ② is a new component to be proposed as P2-02) or 𠃉 二 (TW & HK variant form); suggest creating P2-02 above at §2 and moving this overall component to §3n.<br>This author also suggests using this component as an enclosure component; please see comment on Collins' D13-01 above. |
| D5-02    |   | <b>CJK COMPONENT C-SIMPLIFIED LEARN</b><br>Example: top of 学 / 𦉳 / 𦉴 / 𦉵 / 𦉶 / 𦉷 etc.<br>= CDPO-81DD / UTC-02966<br>Can be decomposed into 𠃉 𠃊, suggest moving to §3n.   |
| D7-01    |   | <b>CJK COMPONENT ARCHITECTURAL STRUCTURE</b><br>Example: top of 亮 / 亭<br>= CDP-8C4D<br>Can be decomposed into 𠃉 𠃊 𠃋, suggest moving to §3n.  |
| D7-02    | <br>   | <b>CJK COMPONENT GUEST</b><br>Example: top of 賓 / 賓, top left of 𦉵<br>= CDP-8D74<br>Decomposable into 𠃉 𠃊 一 少 (𠃉 𠃊 一 少 is a variant), suggest moving to §3n. The bottom subcomponent can be also spun off below as a new, decomposable component H4-01 that could also be added to §3n.  |

**Comment [A9]:** New variant glyph to be used as the enclosing component in 樂 and others.

**Comment [A10]:** Variant of the enclosing component form added above (sometimes used in TW and HK, as in 尾).


**Comment [A11]:** Added Japanese variant glyph for sake of completeness.

**3.1. FS = 1 (FS Code = H)**

| IDS code | Glyph   | Decomposition  | Description   |
|----------|---|----------------|---|
| H4-01    |  | 𠃉 一 少<br>𠃉 一 少 | <b>CJK COMPONENT PART GUEST</b><br>Example: middle of 賓 / 賓, right half of 𦉵, bottom of 𦉶, bottom left of 𦉷, middle of 𦉸<br>New component proposed by this author |

This author also suggests synthesizing a new enclosing component, P4-07, by combining together P2-01 and Z2-01 which are often (but not exclusively) used in pairs. P4-07 should be put into §3n.

**2.3. FS = 3 (FS Code = P)**

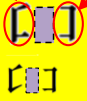
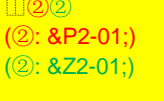
| IDS Code | Glyph   | Description + additions by Vargas                                 |
|----------|---|---|
| P2-01    |  | CJK COMPONENT LEFT ENCLOSURE<br>Example: left of 印, top left of 兜 |

**Comment [A12]:** Added variant glyph for sake of completeness.

**2.5. FS = 5 (FS Code = Z)**

| IDS Code | Glyph   | Description   |
|----------|---|---|
| Z2-01    |  | CJK COMPONENT RIGHT ENCLOSURE<br>Example: right of 兜 / top right of 段 |

**3.3. FS = 3 (FS Code = P)**

| IDS code | Glyph   | Decomposition   | Description   |
|----------|---|---|---|
| P4-07    |  | <br>(②: &P2-01;)<br>(②: &Z2-01;) | CJK COMPONENT ENCLOSURE<br>Example: top enclosure of 兜 / 兜<br>New component proposed by this author |

**Additional new components proposed by this author:**

In addition to the new components proposed above (as a response to comments by SAT, proposals by Collins, and as the result of further analysis of some of the original components proposed by Chan & Lu), this author adds below a few more new components taken from other encoded ideographs, including two decomposable components:


**2.2. FS = 2 (FS Code = S)**

| IDS Code | Glyph   | Description   |
|----------|---|---|
| S7-04    |  | CJK COMPONENT YELLOW LOAM<br>Example: bottom of 董 / 董 |

**2.3. FS = 3 (FS Code = P)**

| IDS Code | Glyph   | Description                                  |
|----------|---|--|
| P7-04    |  | CJK COMPONENT 20137 TOP<br>Example: top of 笄 |

**2.5. FS = 5 (FS Code = Z)**

| IDS Code | Glyph   | Description   |
|----------|---|---|
| Z3-04    |  | CJK COMPONENT DREAD<br>Example: bottom of 喪 / 畏<br>Note: similar to 衣 |

**3.1. FS = 1 (FS Code = H)**

| IDS code | Glyph | Decomposition | Description   |
|----------|-------|---------------|---|
| H5-02    | 吉     | 𠄎𠄎            | CJK COMPONENT SHELL ONE<br>Example: top of 壳, top left of 殼 / 穀<br>= NUCV-264a form 1 |
| H6-04    | 吉     | 𠄎𠄎一           | CJK COMPONENT SHELL TWO<br>Example: top of 壳, top left of 殼 / 穀<br>= NUCV-264a form 2 |

This author also adds to §4 (formerly §3o) an additional component to be provisionally used “for reference”, until UTC-00984 (sequence number 00470 in IRG Working Set 2015) gets encoded into the UCS:

| Base Character | Variant Glyph | Descriptions                                       |
|----------------|---------------|--|
| U+9FB4<br>𠄎    | 𠄎             | Example: bottom of 吉 / 令 (J-source)<br>= UTC-00984 |

**Comment [A13]:** If UTC-00984 from Working Set 2015 gets encoded, consider changing or dropping this component from this list.

## Consolidation of all comments & additions above:

This author, as hinted at the beginning of this feedback document, suggests creating a new list of unencoded decomposable components for IDS use, in order to accommodate many of the components questioned by SAT and most of the new additions suggested by Collins, along with relocated original components that may fit into this category (based on their decomposition) and **most of this author's own proposed additions**.

As stated in the first page of this document, IRG N2225's current section "2. List of Components" is suggested to be renamed to section "2. List of **Atomic Components**" so that it only lists unencoded components that cannot be resolved into IDS sequences of simpler subcomponents; some of SAT's own atomic (and quasi-atomic) component suggestions may also be added to this renamed list, along with **some additions suggested by this author**. Old section "3. Additional Components for Reference" is to be pushed down as renumbered section "4. Additional Components for Reference" in order to accommodate the new proposed list; in addition to its currently-contained variants, one of the Collins suggestions questioned above fits into this field.

This author's new list becomes new section "**3. List of Non-Atomic Components**", and it accommodates unencoded components that can be decomposed into other subcomponents. §3n is to follow the same layout as §2: the list is sub-partitioned according to first stroke. In order to avoid changing components' IDS codes, components from §2 that may be moved into this new section will keep the same IDS code, analogous to Unifiable-Component-Variant (UCV) & Non-Unifiable-Component-Variant (NUCV) rules in IRG Working Document Series (IWDS) documents N2242 & N2243 that maintain their rule numbers even if they switch from one list to another. **Decompositions** are shown in a **new column** inserted between the glyph and description columns.

In order to better visualize the new section, the author has decided to consolidate N2225's original tables with the suggested additions (Collins, SAT, Vargas) and relocations in the following pages:

### 1. Introduction

| FS | FS Code  | Name of Stroke                                  | Rep Glyph | Character Examples         |
|----|----------|---|-----------|----------------------------|
| 1  | <b>H</b> | Horizontal stroke (“橫”, <i>Héng</i> )           | 一         | 一 (one), 三 (three)         |
| 2  | <b>S</b> | Vertical stroke (“豎”, <i>Shù</i> )              | 丨         | 卜 (divination), 巾 (thirty) |
| 3  | <b>P</b> | Left slant stroke (“撇”, <i>Piě</i> )            | 丿         | 川 (river), 月 (moon)        |
| 4  | <b>D</b> | Dot stroke (“點”, <i>Diǎn</i> )                  | 丶         | 流 (current), 宝 (precious)  |
| 5  | <b>Z</b> | Turned stroke (“折”, <i>Zhé</i> ) <sup>1</sup> . | ㇇ ㇈       | 弩, 予, 細                    |

<sup>1</sup> There are many different forms of the turned strokes and their starting stroke can be any of the above single strokes. No distinction is made in this document (IRG N2225).



## 2. List of Atomic Components

### 2.1. FS = 1 (FS Code = H)

| IDS code | Glyph  | Description  |
|----------|--------|--|
| H2-01    | 与      | CJK COMPONENT SIMPLIFIED AND<br>Example: enclosure of 与, middle enclosure of 烏   |
| H3-01    | 𠂇<br>𠂈 | CJK RADICAL PAW FOUR <sup>2</sup><br>Example: bottom of 虐 / 疒, bottom of 虐 / 倉<br>= CDP-8DC1 / UTC-00969<br>(variant glyph = USAT-05803; may be confusable with P3-01 below) |
| H3-02    | 戈      | CJK COMPONENT C-SIMPLIFIED TRIPLE EARTH<br>Example: 尧<br>= CDP-89D9  |
| H3-03    | 𠂇      | CJK COMPONENT CAI<br>Example: outside of 在 / 存<br>= CDP-88F1   |
| H5-01    | 市      | CJK COMPONENT SACK DRAWSTRING WITH STOPPER<br>Example: top of 囊 and 囊<br>= CDP-8DDD  |
| H6-03    | 馬      | CJK COMPONENT PART HORSE<br>Example: outside of 馬<br>= CDP-896A  |
| H7-01    | 𠂇      | CJK COMPONENT LIGATURE DOUBLE EAST<br>Example: top of 曹 / 曹<br>= CDP-8D61  |
| H8-01    | 𠂇      | CJK COMPONENT LEAF<br>Example: top of 𠂇 / 憂 / 憂<br>= CDP-8CD4  |

Comment [A14]: Added variant glyph identical to USAT-05803; could be confusable with P3-01 below.

### 2.2. FS = 2 (FS Code = S)

| IDS Code | Glyph | Description  |
|----------|-------|--|
| S2-01    | 𠂇     | CJK COMPONENT C-SIMPLIFIED MOUND<br>Example: left of 帅 / 归<br>= CDP-8968     |
| S3-01    | 𠂇     | CJK COMPONENT ANCIENT PAIR<br>Example: enclosure of 𠂇, bottom enclosure of 𠂇 |
| S4-02    | 𠂇     | CJK COMPONENT BOTTOMLESS HOLLOW<br>Example: bottom of 𠂇<br>= CDP-8DDF        |
| S5-02    | 𠂇     | CJK COMPONENT REPTILE BODY<br>Example: body of 𠂇, 𠂇<br>= CDP-8BE8            |

<sup>2</sup> The name of the components used in this document (IRG N2225) is only meant as a reference. It has no significance in the IDS sequence. So, its appropriateness will not be discussed for WS 2017. However, if the naming of these components is changed as a result of IRG/WG discussion, update to this document (IRG N2225) will be done accordingly.

|       |   |  |
|-------|---|--|
| S7-02 | 亞 | CJK COMPONENT VASE BOTTOM<br>Example: bottom of 壺 / 壺 / 壺 / 歎                        |
| S7-03 | 甞 | CJK COMPONENT 2E5D3 TOP<br>Example: top of 甞   |
| S7-04 | 堊 | CJK COMPONENT YELLOW LOAM<br>Example: bottom of 堊 / 堊                                |
| S8-01 | 龍 | CJK COMPONENT DRAGON BODY<br>Example: right of 龍<br>= CDP-89B0                       |
| S8-04 | 甞 | CJK COMPONENT RAISE THREE<br>Example: right of 甞; might be unifiable with U+2D048 甞? |
| S9-01 | 𩇛 | CJK RADICAL J-SIMPLIFIED FROG<br>Example: right of 繩 / 蠅                             |

**Comment [A15]:** If this one turns out to be unifiable with 甞, then please move it to §4 below as an "additional component for reference" with U+2D048 甞 as base character.

**2.3. FS = 3 (FS Code = P)**

| IDS Code | Glyph | Description   |
|----------|-------|---|
| P2-01    | 匚     | CJK COMPONENT LEFT ENCLOSURE<br>Example: left of 印, top left of 兜                         |
| P2-02    | 𠃉     | CJK COMPONENT OMISSION RIGHT HALF<br>Example: top right of 樂 / 甞                          |
| P3-01    | 𠃉     | CJK RADICAL PAW THREE<br>Example: left of 印 / 衰 / 衰 / 衰 / 衰 / 歲                           |
| P3-03    | 𠃉     | CJK COMPONENT SINGLE VARIANT ONE<br>Example: right of 𠃉<br>Derived from 个, not 4E87       |
| P3-04    | 𠃉     | CJK COMPONENT SIMPLIFIED PART OVERSIGHT<br>Example: top right corner of 監 / 覽             |
| P3-05    | 𠃉     | CJK COMPONENT PART LIU<br>Example: top left of 留  |
| P4-01    | 𠃉     | CJK RADICAL V-SIMPLIFIED WIND<br>Example: right of 𠃉                                      |
| P4-02    | 夕     | CJK RADICAL MEAT TWO<br>Example: top of 夕 / 蚤 / 然 / 將                                     |
| P4-03    | 夕     | CJK RADICAL MOON TWO<br>Example: top right of 望<br>= CDP-8B5E                             |
| P4-04    | 𠃉     | CJK COMPONENT LIGATURE DOUBLE HUMAN<br>Example: bottom right of 旅 (J-source) / right of 丞 |

**Comment [A16]:** Added variant glyph for sake of completeness.

**Comment [A17]:** Vietnamese tends to use 2 horizontal strokes as in 二

|        |   |  |
|--------|---|--|
| P4-05  | 𪇑 | CJK COMPONENT SIMPLIFIED PART BIRD<br>Example: outside of 鸟 / 島              |
| P5-02  | 𪇒 | CJK COMPONENT LOOSE ROCK ONE<br>Example: left of 段<br>= CDP-8CE4             |
| P6-02  | 𪇓 | CJK COMPONENT BEAST HEAD ONE<br>Example: top of 象 / 象                        |
| P7-01  | 𪇔 | CJK COMPONENT FOOD VESSEL TWO<br>Example: left of 卽(+VS19:卽) / 既(+VS20:既)    |
| P7-02  | 𪇕 | CJK COMPONENT BEAST BODY THREE<br>Example: bottom of 𪇕 / 𪇕 / 𪇕 / 𪇕 / 𪇕 / 𪇕   |
| P7-03  | 𪇖 | CJK COMPONENT PART BIRD<br>Example: outside of 鳥 / 鳥<br>= CDP-8CBB           |
| P7-04  | 𪇗 | CJK COMPONENT 20137 TOP<br>Example: top of 𪇗                                 |
| P8-01  | 𪇘 | CJK COMPONENT DANCE<br>Example: top of 無 / 舞<br>= CDP-8BBE                   |
| P8-03  | 𪇙 | CJK COMPONENT PART OVERSIGHT<br>Example: top right of 𪇙 / 𪇙                  |
| P9-01  | 𪇚 | CJK COMPONENT PART POURING VESSEL<br>Example: top enclosure part of 𪇚 / 𪇚    |
| P10-01 | 𪇛 | CJK COMPONENT RAISE TWO<br>Example: enclosure of 與 / 與 / 與 / 與<br>= CDP-8BA8 |

**Comment [A18]:** Collins had added IVS variation selectors in his feedback, but old machines do not support IVS, so include here bitmap glyphs just in case.

2.4. FS = 4 (FS Code = D)

| IDS Code | Glyph | Description   |
|----------|-------|---|
| D4-02    | 𪇜     | CJK COMPONENT MOVE ON VARIANT ONE<br>Example: bottom right of 𪇜   |
| D5-03    | 𪇝     | CJK COMPONENT MOVE ON VARIANT TWO<br>Example: right of 𪇝  |
| D6-01    | 𪇞     | CJK COMPONENT C-SIMPLIFIED CHAOS<br>Example: top of 恋 / 𪇞 / 𪇞.<br>Note: use 4EA6 亦 instead where appropriate, such as 5955 𪇞<br>= UTC-02967 |

**Comment [A19]:** Source of this ideograph? I'd appreciate it if sources were included here for Collins' three unencoded example ideographs.

**Comment [A20]:** Ditto.

|        |   |  |
|--------|---|--|
| D6-02  | 良 | CJK COMPONENT VIRTUOUS<br>Example: left of 朗 / 朗, left hand side form of 良<br>= CDP-8C42 |
| D6-03  | 廌 | CJK COMPONENT BEAST HEAD TWO<br>Example: outside of 鹿 / 鹿<br>= CDP-8D56                  |
| D10-01 | 廾 | CJK COMPONENT EVEN<br>Example: top of 齊 / 齊<br>= CDP-8DEB                                |

2.5. FS = 5 (FS Code = Z)

| IDS Code           | Glyph  | Description   |
|--------------------|--------|---|
| Z2-01              | 勹      | CJK COMPONENT RIGHT ENCLOSURE<br>Example: right of 兜 / top right of 段   |
| Z2-02              | 勹      | CJK COMPONENT LEFT HAND ONE<br>Example: left of 収   |
| Z2-03              | 匕      | CJK COMPONENT LEFT HAND TWO<br>Example: right of 𠂇<br>= CDP-88B5  |
| Z2-04              | 𠂇      | CJK COMPONENT SIMPLIFIED PART HORSE<br>Example: outside of 马  |
| Z2-05              | 𠂇      | CJK COMPONENT RIGHT HAND TWO<br>Example: right top of 祭 / 登<br>= CDP-8C4E   |
| Z3-01              | 𠂇<br>𠂇 | CJK COMPONENT RIGHT HAND ONE<br>Example: top of 尋 / top right of 侵 / bottom of 彗<br>= CDP-8BAB                                      |
| Z3-02              | 𠂇      | CJK COMPONENT HOLD ONE<br>Example: right top of 恐 / 鞏 etc.  |
| Z3-03 <sup>3</sup> | 丸      | CJK COMPONENT HOLD TWO (variant of 𠂇)<br>Example: right of 執 / 執 / 狽 / 秋<br>Note: use 4E38 丸 (wan2) where appropriate, e.g. 7D08 紉. |
| Z3-04              | 𠂇      | <b>CJK COMPONENT DREAD</b><br>Example: bottom of 喪 / 畏<br>Note: similar to 衣  |
| Z4-01              | 𠂇      | CJK COMPONENT TILTED HEAD<br>Example: bottom of 吳<br>= CDP-8CD5   |
| Z4-02              | 𠂇      | CJK COMPONENT RAN<br>Example: left of 那<br>= CDP-89FA   |

Comment [A21]: Added variant glyph for sake of completeness.

<sup>3</sup> Refer to IRG N2225 Appendix A: Rationale for Inclusion of Z3-03 (reprinted in this feedback just in case).

|       |   |  |
|-------|---|--|
| Z5-01 | 艮 | CJK COMPONENT FOOD VESSEL ONE<br>Example: left of 即 / 既<br>= CDP-8B7C                          |
| Z5-02 | 𠃉 | CJK COMPONENT LOOSE ROCK TWO<br>Example: left of 段<br>= CDP-8C7A                               |
| Z7-01 | 鼠 | CJK COMPONENT BEAST BODY THREE<br>Example: bottom of 鼠 / 𧰨 (TW/HK variant glyph)<br>= CDP-8D46 |

### 3. List of Non-Atomic Components

#### 3.1. FS = 1 (FS Code = H)

| IDS code | Glyph  | Decomposition | Description   |
|----------|--------|---------------|---|
| H3-04    | 六      | 𠃉一八           | CJK COMPONENT RAISE ONE<br>Example: bottom of 具 / 巽 / 共   |
| H4-01    | 彳      | 𠃉一少<br>𠃉一少    | CJK COMPONENT PART GUEST<br>Example: middle of 賓 / 賓, right half of 杪, bottom of 夢, bottom left of 頰, middle of 闌 |
| H4-02    | 土      | 𠃉十            | CJK COMPONENT COMET TOP<br>Example: top of 索, 宇, 寔 etc.<br>= CDP-8D52   |
| H5-02    | 吉      | 𠃉土            | CJK COMPONENT SHELL ONE<br>Example: top of 壳, top left of 殼 / 殼<br>= NUCV-264a form 1                             |
| H6-01    | 廌      | 𠃉&H2-01;...   | CJK COMPONENT BEAST BODY TWO<br>Example: bottom of 廌 / 廌 / 廌<br>= CDP-8DBF  |
| H6-02    | 𠃉      | 𠃉一𠃉一          | CJK COMPONENT LIGATURE DOUBLE ARRIVE<br>Example: top of 晉<br>= CDP-8DAB   |
| H6-04    | 吉      | 𠃉土            | CJK COMPONENT SHELL TWO<br>Example: top of 壳, top left of 殼 / 殼<br>= NUCV-264a form 2                             |
| H9-01    | 𠃉<br>𠃉 | 𠃉𠃉&P3-04;     | CJK COMPONENT SIMPLIFIED OVERSIGHT<br>Example: top of 監 / 覽 / 𠃉<br>= CDP-8CC9                                     |
| H9-03    | 𠃉      | 𠃉𠃉衣           | CJK COMPONENT HELP VARIANT ONE<br>Example: bottom right of 爆  |
| H14-01   | 臨      | 𠃉𠃉&P8-03;     | CJK COMPONENT OVERSIGHT<br>Example: top of 鑿 / 覽<br>= CDP-8D50  |
| H14-02   | 𠃉      | 𠃉𠃉𠃉米八女        | CJK COMPONENT ZONG VARIANT ONE<br>Example: right of 𠃉   |

**Comment [A23]:** Vietnamese prefers 𠃉, not 𠃉, should that be a separate component?

**Comment [A22]:** Added variant glyph for Vietnamese, in response to Collins' question above.

**3.2. FS = 2 (FS Code = S)**

| IDS code | Glyph | Decomposition                          | Description   |
|----------|-------|--|---|
| S4-01    | 収     | 𠄎&S2-01;又                              | CJK COMPONENT C-SIMPLIFIED LEAD<br>Example: top of 堅 / 豎 / 緊 / 腎 / 賢<br>= CDPO-81EF / UTC-02964 |
| S4-03    | 𠄎     | 𠄎                                      | CJK COMPONENT 2A7FD VARIANT<br>Example: right of 𠄎  |
| S5-01    | 𠄎     | 𠄎&S2-01;&P3-04;                        | CJK COMPONENT C-SIMPLIFIED OVERSIGHT<br>Example: top of 監 / 覽<br>= CDPO-81F0 / UTC-02965        |
| S7-01    | 𠄎     | 𠄎&S3-01;<br>𠄎&S3-01;                   | CJK COMPONENT COCOON<br>Example: outer enclosure of 繭 / 繭                                       |
| S8-02    | 𠄎     | 𠄎                                      | CJK COMPONENT GAUZE VARIANT ONE<br>Example: left of 𠄎, right of 𠄎                               |
| S8-03    | 𠄎     | 𠄎                                      | CJK COMPONENT GAUZE VARIANT TWO<br>Example: right of 𠄎, bottom of 𠄎                             |
| S11-01   | 𠄎     | 𠄎<br>𠄎                                 | CJK COMPONENT DREAMING PERSON<br>Example: top of 夢 / 薨 / 蓍 / 蓍 / 蓍 / 蓍<br>= CDP-8D60            |
| S11-02   | 𠄎     | 𠄎③<br>(②: &S8-04; 𠄎)<br>(③: UTC-01671) | CJK COMPONENT RAISE VARIANT<br>Example: right of 𠄎  |

**3.3. FS = 3 (FS Code = P)**

| IDS code | Glyph | Decomposition                       | Description  |
|----------|-------|-------------------------------------|--|
| P4-06    | 𠄎     | 𠄎                                   | CJK COMPONENT ORDER RIGHT HALF<br>Example: right of 𠄎                |
| P4-07    | 𠄎     | 𠄎②②<br>(②: &P2-01;)<br>(②: &Z2-01;) | CJK COMPONENT ENCLOSURE<br>Example: top enclosure of 兜 / 兜           |
| P5-01    | 𠄎     | 𠄎&P3-05;刀                           | CJK COMPONENT LIU<br>Example: top of 留 / top left of 劉<br>= CDP-8C69 |
| P6-01    | 𠄎     | 𠄎                                   | CJK COMPONENT BEAST BODY ONE<br>Example: bottom of 𠄎<br>= CDP-8C40   |

|        |             |           |  |
|--------|-------------|-----------|--|
| P6-03  | 𠂇<br>𠂈<br>𠂉 | 𠂇𠂈𠂉       | CJK COMPONENT COMPLEX PERSON CAVE<br>Example: top of 𠂇 / 𠂈 / 𠂉<br>= CDP-8CAF |
| P8-02  | 𠂊           | 𠂊         | CJK COMPONENT LADLE VARIANT<br>Example: right of 𠂊                           |
| P15-01 | 𠂋<br>𠂌<br>𠂍 | 𠂋&P9-01;同 | CJK COMPONENT COMPLEX POURING VESSEL<br>Example: top of 𠂋 / 𠂌                |

**3.4. FS = 4 (FS Code = D)**

| IDS code | Glyph       | Decomposition                        | Description   |
|----------|-------------|--------------------------------------|---|
| D4-01    | 𠂎<br>𠂏<br>𠂐 | 𠂎 &P2-02;<br>𠂏 二二 (TW/HK)            | CJK COMPONENT OMISSION<br>Example: bottom of 𠂎 / 𠂏, enclosure of 𠂐 / 𠂑 / 𠂒(尾)<br>= CDP-89AE               |
| D4-03    | 𠂑           | 𠂑                                    | CJK COMPONENT ANCIENT FAULT TOP<br>Example: top of 𠂑, 帝, 商, 商, 产 etc.<br>= CDP-8BAE                       |
| D5-02    | 𠂒           | 𠂒                                    | CJK COMPONENT C-SIMPLIFIED LEARN<br>Example: top of 学 / 𠂒 / 𠂓 / 𠂔 / 𠂕 / 𠂖 etc.<br>= CDPO-81DD / UTC-02966 |
| D7-01    | 𠂗           | 𠂗                                    | CJK COMPONENT ARCHITECTURAL STRUCTURE<br>Example: top of 亮 / 亭<br>= CDP-8C4D                              |
| D7-02    | 𠂘<br>𠂙      | 𠂘&H4-01;                             | CJK COMPONENT GUEST<br>Example: top of 𠂘 / 𠂙, top left of 𠂚<br>= CDP-8D74                                 |
| D8-01    | 𠂛           | 𠂛                                    | CJK COMPONENT RIGHTEOUS VARIANT ONE<br>Example: right of 𠂛, 𠂜, 𠂝  |
| D9-01    | 𠂞           | 𠂞                                    | CJK COMPONENT TRIGRAM FIVE VARIANT ONE<br>Example: left of 𠂞  |
| D13-01   | 𠂟           | 𠂟④同③<br>(④: &D4-01;)<br>(③: &H3-04;) | CJK COMPONENT PROSPER VARIANT ONE<br>Example: left of 𠂟, bottom of 𠂟                                      |

**Comment [A24]:** New variant glyph to be used as enclosing component in 𠂐 and others.

**Comment [A25]:** Variant of the enclosing component form added above (sometimes used in TW and HK as in 尾).

**Comment [A26]:** Added Japanese variant glyph for sake of completeness.

**Comment [A27]:** Source of this ideograph?

**3.5. FS = 5 (FS Code = Z) (currently none allocated here)**

#### 4. Additional Components for Reference

| Base Character | Variant Glyph | Descriptions   |
|----------------|---------------|--|
| U+9FB4<br>𠄎    | 𠄎             | Example: bottom of 𠄎 / 令 (J-source)<br>= UTC-00984         |
| U+2B94E<br>𠄎   | 𠄎             | Example: right of 𠄎 / 𠄎                                    |
| U+53C0<br>𠄎    | 𠄎             | Example: top of 𠄎 / 𠄎<br>= CDP-8BD0                        |
| U+26C29<br>𠄎   | 𠄎             | Example: right of 𠄎  |
| U+5C6F<br>𠄎    | 𠄎             | Example: right of 𠄎 / 𠄎 / 𠄎<br>= CDP-8AF0                  |
| U+6B62<br>止    | 𠄎             | Example: bottom of 𠄎 / 𠄎<br>= CDP-8962                     |
| U+9FB0<br>𠄎    | 𠄎             | Example: bottom of 𠄎 / 𠄎 / 𠄎<br>= CDP-8961                 |
| U+3B30<br>𠄎    | 𠄎             | Example: right of 𠄎 / 𠄎 / 𠄎<br>= CDP-89E7                  |
| U+4E23<br>𠄎    | 𠄎             | Example: top right of 𠄎, right of 𠄎 / 𠄎 / 𠄎<br>= CDPO-85FD |

**Comment [A28]:** If UTC-00984 from WS 2015 gets encoded, consider changing or dropping this.

**Comment [A29]:** Formerly H11-01 by Collins

**Comment [A30]:** Cropped padding and stretched glyph so as to make it consistent with all other glyphs in this list

#### 5. Questionable Components for Further Discussion

| IDS code | Glyph | Description  |
|----------|-------|--|
| H9-02    | 𠄎     | <b>CJK COMPONENT 78CA VARIANT ONE</b><br>The author of this feedback cannot see how to fit this Collins-proposed component (which was already submitted to IRG Working Set 2015 as a standalone ideograph with sequence num. 02833, source UTC-01039, and decomposition 𠄎石双) anywhere within the consolidated lists above; to further complicate things, no encoded or WS2015-submitted ideograph uses this as a component (other than UTC-01039 itself). This author invites further discussion on how to deal with this component (whether to add it to §3n or to §4 above, or anything else?) |



## Appendix A: Rationale for Inclusion of Z3-03

Z3-03 丸 is included as a separate component from U+4E38 丸. First, they are considered different components in the Taiwan and Hong Kong locale. For characters whose etymology is U+4E2E 𠂔 (to hold), the glyph shape Z3-03 with the dot protruding is used. For characters whose etymology is U+4E38 丸 (small round object), the T/H glyph shape of U+4E38 with the dot not protruding is used.

This is a similar case to the existing dis-unified radicals 2EC0 𠂔 CJK RADICAL GRASS THREE from 2EBE 𠂔 CJK RADICAL GRASS ONE, and 2EBC 月 CJK RADICAL MEAT from 2F49 月 KANGXI RADICAL MOON. The difference in the shape is reflective of a different etymology and hence “non-cognateness” of characters in the Taiwan and the Hong Kong locales.

Second, these normally unifiable differences in components could also reflect a difference in meaning in certain context for other locales, and not just as a simple typographical variation. For example, PRC selectively applies such typographical variations as a standardized difference for a range of non-cognate characters. The difference between U+2EC0(𠂔) and U+2EBE(𠂔) is used to distinguish between the two characters U+830D 苟 and U+82DF 苟 in the PRC locale.

Thus, the inclusion of Z3-03 component is also meaningful and useful for IDS composition.

### Final thoughts:

Although the current immediate goal of Chan & Lu's IRG N2225 is to possibly establish a new Working Document Series (WDS), this author believes that, should these new components proposed above prove to be very useful in optimizing the IDS sequencing of currently-encoded CJK ideographs and future working-set and urgently-needed character submissions, these components should not just become a brand-new WDS, but also be proposed to WG2 as a new collection of CJK-related characters to be encoded into both the 13 unassigned codepoints of the **CJK Radicals Supplement** block (U+2E80 – U+2EFF) and a brand-new block suggested by Chan's IRG N2204 to be named **CJK Radicals Extended-A**. This author believes that such block be renamed to **CJK Components Supplement**, because none of these components proposed above (save for the six components with RADICAL in their provisional names, that is, H3-01, S9-01, P3-01, and P4-01 thru P4-03) are used as radicals but rather as phonetic components (or elements of phonetic components); also, this author suggests using up the empty 16-codepoint range **U+2FE0 – U+2FEF** (between the **Kangxi Radicals** [U+2F00 – U+2FDF] & **Ideographic Description Characters** [U+2FF0 – U+2FFF] blocks) to encode the 16 most commonly-used components among those listed above.

So, this author suggests that the IRG Rapporteur propose to WG2 the six unencoded radicals H3-01, S9-01, P3-01, and P4-01 thru P4-03 above to be encoded into any of the 13 free codepoints of **CJK Radicals Supplement** and then, depending on the usefulness of the remaining unencoded components in IRG's current and future IDS work, revise the consolidated lists above and propose to WG2 the creation of two new blocks: **CJK Components** (U+2FE0 – U+2FEF), where the 16 most frequently encountered of all components proposed above can be allocated, and **CJK Components Supplement** (it could be allotted into either the SMP, the SIP, or possibly the TIP, because there is no room left in the BMP for a new CJK-related block other than the **CJK Components** block suggested by this author), where the remaining components accepted from the lists above can be encoded. (It is worth noting that some of the components listed above have been also submitted to IRG Working Set 2015 as standalone ideographs, so whether they are considered for "duplicate" encoding as components or not will depend on how early they are encoded as ideographs into the future **CJK Unified Ideographs Extended-G** block.)

Since it has been previously proposed by the Unicode Consortium in IRG N2117 to create an IDS property into the Unihan Database (this author suggests making such property an informational one), the components proposed above, if accepted for encoding into the UCS, could be used to help populate this recently-proposed Unihan property. These components could also be incorporated into the U-source database ([USourceData.txt](#)) established as part of Unicode Standard Annex #45 (UAX45), which currently uses the "?" placeholder for indescribable, non-decomposable unencoded components.

# Feedback on IRGN2225 “CJK Supplementary Components for IDS Use”

Andrew West

2017-10-02

1. IRGN2225 states that the IDS code is preceded by &; (e.g. "&;H7-01"). I think this is a mistake, and the document should state that the ampersand should precede the IDS code, and the semicolon should terminate the IDS code (e.g. "&H7-01;").
2. Remove H3-02 戈 as it is the same as U+2D80D 戈.
3. Consider removing S7-01 苜 as it only occurs in one character (繭), which can be represented as 𠂇𠂈𠂉𠂊𠂋𠂌 𠂍 | 虫.
4. Add separate components for two forms of S7-02.
5. Add another component for PRC form of left side of U+537D 卽 (i.e. 卩) as well as P7-01.
6. Add separate components for two forms of P7-02.
7. Add separate components for three forms of P15-01.
8. Add top of 祭/登 as a separate component instead of representing it as a combination of P4-02/P4-03 and Z2-05.
9. Remove Z4-02 月 as it is the same as U+2D0C2 月.
10. Consider adding the following additional components which are used in the BabelStone public IDS file (<http://www.babelstone.co.uk/CJK/IDS.TXT>):
  - top of 專惠 (𠂇) – distinguish from U+ 53C0 𠂇.
  - right of 托 (𠂇) – distinguish from U+4E47 𠂇.
  - right of 竹𠂇 (𠂇) – distinguish from U+4E87 𠂇.
  - left of 𠂇, right of 𠂇 (𠂇)
  - top of 𠂇 (𠂇)
  - right of 浦 (𠂇)
  - right of 𠂇 (𠂇) – distinguish from U+3B30 𠂇.

- top of 藁 (声)
- bottom of 玄 (幺)
- outside of 武貳 (弋)
- outside of 幾畿 (幾)
- inside of 鹵 (久)
- right of 帛 (仄)
- left of 鞠 (拳)
- right of 洸 (晃)
- bottom of 窰 (缶)
- right of 駟 (亞)