

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION  
ORGANISATION INTERNATIONALE DE NORMALISATION  
ISO/IEC JTC 1/SC 2/WG 2/IRG**

**Universal Coded Character Set  
(UCS)**

**ISO/IEC JTC 1/SC 2/WG 2/IRG N2230  
2018-3-19**

<b>Title:</b>	<b>SAT's Submission for the IRG Working Set 2017</b>
<b>Source:</b>	<b>SAT Daizōkyō Text Database Committee (Prof. Masahiro Shimoda)</b>
<b>Action:</b>	<b>To be considered by IRG</b>
<b>Distribution:</b>	<b>IRG Member Bodies and Ideographic Experts</b>

SAT proposes 305 ideographs for the IRG Working Set 2017 in the CJK unified ideographs. The ideographs were collected from Taisho Shinshu Daizokyo (大正新脩大藏經, Tokyo, 1924, henceforth, Taisho) except a character. If a shape of the glyph is not clear, this evidence file indicates relatively clear image of the same glyph from Taisho's witnesses, such as woodcut printings and manuscripts mainly below:

1. 高麗版大藏經(Tripitaka Koreana) second edition,  
京城, 1236-1251.  
Research Institute of Tripitaka Koreana [http://kb.sutra.re.kr/ritk\\_eng/index.do](http://kb.sutra.re.kr/ritk_eng/index.do) (RITK高麗藏)  
National Diet Library, <http://kindai.ndl.go.jp/info:ndljp/pid/1240226> (高麗藏)
2. 大日本校訂縮刷大藏經 (Dainihon Koutei Shukusatsu Daizokyo)  
縮刷大藏經刊行会 (Shukusatsu Daizokyo Kankoukai), 1939 (reprint). (縮刷藏)
3. 頻伽精舍校刊大藏經 (Binga Shouja Koukan Daizokyo)  
頻伽精舍, 1923.
4. 白蓮社刊一切經音義 (Byakurensha Kan Issai Kyo Ongi) (白蓮社)  
白蓮社, 1738.

Appendix 1: Attributes  
[IRGN2230SAT\\_attributes\\_v2.xlsx](#)

Appendix 2: Bitmap files  
[http://bauddha.dhii.jp/sat\\_exth/satws2017\\_bmps\\_v2.zip](http://bauddha.dhii.jp/sat_exth/satws2017_bmps_v2.zip)

Appendix 3: Evidence files  
[http://bauddha.dhii.jp/sat\\_exth/IRGN2230SATWS2017Evidences\\_v2\\_1.pdf](http://bauddha.dhii.jp/sat_exth/IRGN2230SATWS2017Evidences_v2_1.pdf)  
[http://bauddha.dhii.jp/sat\\_exth/IRGN2230SATWS2017Evidences\\_v2\\_2.pdf](http://bauddha.dhii.jp/sat_exth/IRGN2230SATWS2017Evidences_v2_2.pdf)  
[http://bauddha.dhii.jp/sat\\_exth/IRGN2230SATWS2017Evidences\\_v2\\_3.pdf](http://bauddha.dhii.jp/sat_exth/IRGN2230SATWS2017Evidences_v2_3.pdf)  
[http://bauddha.dhii.jp/sat\\_exth/IRGN2230SATWS2017Evidences\\_v2\\_4.pdf](http://bauddha.dhii.jp/sat_exth/IRGN2230SATWS2017Evidences_v2_4.pdf)

Appendix 4: Auxiliary components  
[IRGN2230\\_SATAuxiliaryComponents.pdf](#)

## IRG Repertoire Submission Summary Form

**ISO/IEC JTC 1/SC 2/WG 2/IRG  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITION OF CJK UNIFIED IDEOGRAPHS TO THE REPERTOIRE OF ISO/IEC 10646**

**Submitters are reminded to:**

1. Fill in all the sections below.
2. Read the Principles and Procedures Document (P & P) available at <http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg45/IRGN2092PnPv8.pdf> for guidelines and details before filling in this form.
3. Use the latest Form from [http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg45/IRGN2092PnP\\_BlankDataFile.xls](http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg45/IRGN2092PnP_BlankDataFile.xls)

See also <http://appsrv.cse.cuhk.edu.hk/~irg/irgwds.html> for the latest *Unifiable Component Variations*.

### A. Administrative

<b>1. IRG Project Code:</b>	IRG Working Set 2017
<b>2. Title:</b>	SAT's Submission for the IRG Working Set 2017
<b>3. Submitter's Region/Country Name:</b>	SAT Daizōkyō Text Database Committee
<b>4. Submitter Type (National Body/Individual Contribution):</b>	Liaison
<b>5. Submission Date:</b>	2018-03-19
<b>6. Requested Ideograph Type (Unified or Compatibility Ideographs)</b>	Unified Ideographs
If Compatibility, does the submitter have the intention to register them as IVS (See UTS #37) with the IRG's approval? (Registration fee will not be charged if authorized by the IRG.)	N/A
<b>7. Proposal Type (Normal Proposal or Urgently Needed)</b>	Normal Request
<b>8. Choose one of the following:</b>	
This is a complete proposal	X
(or) More information will be provided later.	

### B. Technical – General

1. Number of ideographs in the proposal:	305
2. Glyph format of the proposed ideographs: (128x128 Bitmap files or TrueType font file)	Bitmap
If Bitmap files, are their file names the same as their source references?	Yes
If TrueType font file, are all the proposed glyphs put into BMP PUA area?	N/A
If TrueType font file, are data for source references vs. character codes provided?	N/A
<b>3. Source references:</b>	
Do all the proposed ideographs have a unique, proper source reference (member body/international consortium abbreviation followed by no more than 9 alphanumeric characters)?	Yes
<b>4. Evidence:</b>	
a. Do all the proposed ideographs have a separate evidence document which contains at least one scanned image of printed materials (preferably dictionaries)?	Yes

b. Do all the printed materials used for evidence provide enough information to track them by a third party (ISBN numbers, etc.)?

Yes

5. Attribute Data Format: (Excel file or CSV text)

*Excel*

## C. Technical - Checklist

<b>Understanding of the Unification Principles</b>		
1. Has the submitter read ISO/IEC 10646 Annex S and does the submitter understand the unification principles?		Yes
2. Has the submitter read the “Unifiable Component Variations” (contact the IRG technical editor through the IRG Rapporteur for the latest version) and does the submitter understand the unifiable variation examples?		Yes
3. Has the submitter read the IRG PnP document and does the submitter understand the 5% Rule?		Yes
<b>Character-Glyph Duplication</b> ( <a href="http://www.itscj.ipsj.or.jp/sc2/open/pow.htm">http://www.itscj.ipsj.or.jp/sc2/open/pow.htm</a> contains all the published ones and those under ballot)		
4. Has the submitter checked that the proposed ideographs are <b>not unifiable</b> with any of the unified or compatibility ideographs of the latest version of ISO/IEC 10646?		Yes
If the checking has been done against an earlier version of ISO/IEC 10646, please specify the version? (e.g. 10646:2012)		10646:2014
5. Has the submitter checked that the proposed ideographs are <b>not unifiable</b> with any of the ideographs in the amendments, if any, of the latest version of ISO/IEC 10646?		N/A
If yes, which amendment(s) has the submitter checked?		N/A
6. Has the submitter checked that the proposed ideographs are <b>not unifiable</b> with any of the ideographs in the proposed amendments, if any, of ISO/IEC 10646?		Yes
If yes, which draft amendment(s) has the submitter checked?		Extension F
7. Has the submitter checked that the proposed ideographs are <b>not unifiable</b> with any of the ideographs in the current working M-set and D-set of the IRG? (Contact IRG chief editor and technical editor through the IRG Rapporteur for the newest list)		Yes
If yes, which document(s) has the submitter checked?		IRG N2223
8. Has the submitter checked that the proposed ideographs are <b>not unifiable</b> with any of the over-unified or mis-unified ideographs in ISO/IEC 10646? (See Annex E of the IRG PnP document).		Yes
9. Has the submitter checked whether the proposed ideographs have any <b>similar ideographs</b> in the current standardized or working sets mentioned above?		Yes
10. Has the submitter checked whether the proposed ideographs have any <b>variant ideographs</b> in the current standardized or working sets mentioned above?		Yes
<b>Attribute Data</b>		
11. Do all the proposed ideographs have attribute data such as the Kangxi radical code and stroke count?		Yes
12. Are there any simplified ideographs (ideographs that are based on the policy described in 簡化字總表) among the proposed ideographs?		No
If yes, does the proposal include proper simplified/traditional indication flag for each proposed ideograph in the attribute data?		N/A
13. Do all the proposed ideographs have the document page number of evidence documents in the attribute data?		Yes
14. Do all the proposed ideographs have the proper Ideographic Description Sequence (IDS) in the attribute data?		Yes
If no, how many proposed ideographs do not have the IDS?		N/A

15. If the answer to question 9 or 10 is yes, do the attribute data include any information on similar/variant ideographs for the proposed ideographs?
16. Do all the proposed ideographs contains the total stroke count(kTotalStrokes)?

Yes

Yes

# SAT Auxiliary Components for IDS Use

## 1. Introduction

This document provides the list of auxiliary components which are and will be found in SAT proposals for IRG Working Sets (WS2017 and thereafter). These components are currently unencoded in UCS, but intended to serve to improve integrity and quality of the IDS-based checking process during deliberation by IRG experts.

The document will be updated as necessary.


## 2. Nomenclature

Components are named with a convention similar to that of IRGN2225, namely:

SAT-**X**YY**Z**

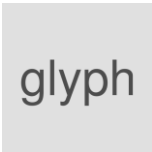
Where,

- **X** is a Latin letter **H, S, P, D, or Z** which denotes the first stroke type of the component. Each stands for the first stroke type 1 (**Héng**), 2 (**Shù**), 3 (**Piě**), 4 (**Diǎn**), and 5 (**Zhé**) defined in the IRG PnP Annex K, respectively.
- **YY** is a two-digit number denotes the total stroke count of the component.
- **Z** is a Latin letter **A-Z** that identifies each unique component from those have identical **XY**. The letter will be assigned from A to Z in the order of our recognition of new component.

The components are represented in IDS with surrounding **&** and **;**, following the convention of XML character entity reference. For example, a character that contains a component SAT-P08A would have an IDS such as 木&SAT-P08A;.

## 3. List

The list of components is ordered according to first stroke, then to the alphanumeric order. Each entry consists of a set of information illustrated below:

SAT-XYZ	<i>Preferred IDS decomposition</i>	<i>Ex. of UCS chars using this component</i>
	<i>Description</i> → <i>Other sources</i>	
(variants)		

Reference names of external sources:

[CHISE]: <http://www.chise.org/ids/>


[GW]: <https://glyphwiki.org/>

### 3.1. Horizontal Bar (H)

<p>SAT-H06A</p> <p>𠄎</p> <p>𠄎</p>	<p>𠄎大干</p> <p>A conflated shape that is related to either 𠄎 (U+3694) or 𠄎 (U+2B889).</p>	<p>𠄎 𠄎</p>
<p>SAT-H08A</p> <p>𠄎</p> <p>𠄎</p>	<p>𠄎牙文</p> <p>One of variants of 𠄎 (U+20A7A) that may occur with or without 𠄎 part.</p>	<p>𠄎 𠄎 𠄎 𠄎</p>
<p>SAT-H08B</p> <p>𠄎</p> <p>𠄎</p>	<p>𠄎士示</p> <p>Variant of 崇 (U+795F) or 𠄎 (U+67F0), seen on the left side of 𠄎 and 𠄎. → CDP-8C6A [CHISE][GW]</p>	<p>𠄎 𠄎 𠄎 𠄎</p>



### 3.2. Vertical Bar (S)


<p>SAT-S04A</p> <p>𠄎</p> <p>𠄎</p>	<p>𠄎𠄎口</p> <p>Typically an alternative development of 𠄎 (U+758B) from Small Seal script.</p>	<p>𠄎 𠄎 𠄎 𠄎</p>
<p>SAT-S07A</p> <p>𠄎</p> <p>𠄎</p> <p>𠄎</p>	<p>𠄎   臣</p> <p>An alternate shape of 臣 (U+268DD). The shape with leftward sweep on the first stroke still comes under this component.</p>	<p>𠄎 𠄎 𠄎 𠄎</p>
<p>SAT-S10A</p> <p>𠄎</p> <p>𠄎</p>	<p>𠄎日共</p> <p>A modern rendering of the top complex of 暴 (U+3B25), typically as the top part of 暴.</p>	<p>𠄎 𠄎 𠄎 𠄎</p>



SAT-S12A	𠄎𠄎大	暴 暴 暴 暴
	The full rendering of the top complex of 暴 (U+3B25), the fuller shape of 暴.	

3.3. Slash (P)

SAT-P03A	𠄎个又 / 𠄎𠄎又	
	An alternative development of 支 (U+652F) from the Seal–Cleric glyph. → IRG WS2015 00511 (USAT09010)	



SAT-P06A	𠄎丿又 / 𠄎𠄎𠄎	𠄎 𠄎 𠄎 𠄎
 	A conflated shape that is related to either 𠄎 (U+53D0) or 𠄎 (U+592D). The shape without linking between 3 <sup>rd</sup> and 4 <sup>th</sup> stroke still comes under this component.	

SAT-P08A	𠄎从开	
	An alternative shape of 𠄎 (U+22199). → IRG WS2015 01117 (USAT08402)	

SAT-P09A	𠄎八入目	睿 睿 睿 睿
 	The bottom complex of 睿 (U+777F). The shape with top two dots flipped still comes under this component.	



### 3.4. Dot (D)

SAT-D06A	𠄎、平	𠄎 𠄎 𠄎 𠄎
 	A cursivized shape of 卒 (U+5352). The shape with middle dots flipped still comes under this component.	

### 3.5. Turn (Z)

Currently does not have entries.

(End of document)