

Subject: A few considerations on CJK Supplementary Components for IDS

Date: 2021.03.08.

Author: KIM, Kyongsok (Republic of Korea)

Status: Individual Contribution

1. Background

Two important IRG documents related with CJK Supplementary Components for IDS are shown below. IRG Recommendation M48.10 (shown in 1.1 below) recommends to produce IRG N2225 (shown in 1.2 below).

1.1 IRGN2220_IRG48RecommendsUpdate.pdf

Recommendation IRG M48.10: CJK Supplementary Components (IRGN2204, IRGN2218)

Unanimous

IRG reviewed the proposal to add CJK components to improve the quality of IDS database. IRG recommends the proposer to work with the IRG Rapporteur to produce an IRG working document for WS 2017 submissions by 2017-07-07(IRGN2225).

1.2 IRGN2225.pdf (CJK Supplementary Components for IDS Use)

Source:	Henry Chan and Qin Lu (IRG Rapporteur)
Title:	CJK Supplementary Components for IDS Use

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:

- About 45 CJK Supplementary Components (will be referred to as "Supp. Comps." hereafter) are listed in IRG N2225.
- For example, in "&H7-01;" (will be referred to as "Supp. Comp. Name" hereafter), H is an FS (first stroke) code, 7 is an SC (stroke count of Supp. Comp.), and 01 is an SN (serial number).
- The objective of Supp. Comps. is "(to) increase quality of machine checking of IDS" (will be referred to as "better machine checking of IDS" hereafter).

2. Some issues to be discussed

Currently, it seems that IRG N2225 is the only guideline in using Supp. Comps. in IDS for WS2017. If Supp. Comps. are to be used in IDS, there seem several issues to be discussed and resolved.

2.1 IRG PnP

- If IRG wants to allow Supp. Comps. to be used in IDS, then IRG should include relevant information in IRG PnP.

2.2 Criteria and procedure to review and accept newly suggested Supp. Comps.

- For example, at the time of WS2017 submission, TCA suggested about 45 NEW Supp. Comps. in appendix 4 of IRG N2231 and SAT suggested about 12 NEW Supp. Comps. whose prefix is "SAT" (not one of H S, P, D, or Z) in IRGN2230_SATAuxiliaryComponents.pdf.

- It seems that IRG did not record in M48 Recommendations whether or not to allow NEW Supp. Comps. in addition to 45 in IRG N2225 to be used in IDS. (Please correct me, if I am wrong.)

- To maintain the list of Supp. Comps. (probably as an IWDS?), IRG need to set up criteria and a procedure to review and accept or reject newly suggested Supp. Comps.

- To avoid confusion, the list of Supp. Comps. need to have a version number and the list need be published at IRG meeting web site.

- IRG PnP probably need to include a form in IRG PnP that MBs (member bodies) can use to suggest additional new Supp. Comps.

2.3 Supp. Comp. Naming procedure

- It seems that MBs can create and use new Supp. Comps. in WS submission.

- If new Supp. Comps. are left as their initial submissions without reviewing, there are two possible problems.

. Two MBs could create and use two distinct Supp. Comps. with the same Supp. Comp. Name.

. Two MBs could create and use two distinct Supp. Comps. Names for the same Supp. Comp.

- To prevent such confusion, IRG need to set up some procedure to prevent collision of Supp. Comp. Names and to ensure a unique Supp. Comp. name for each Supp. Comp.

- In IRG N2225, the FS code can be one of H S, P, D, or Z.

- IRG need to decide whether or not FS code other than those five in IRG N2225 is allowed.

- For example, SAT used "SAT" as FS code (?) as in "&SAT-H06A;" and Viet Nam used "CDP" as in "&CDP-89DF;" in WS2017 submissions. Supp. Comps. with FS "SAT" or "CDP" remain in IRG N2444, IRG Working Set 2017 Version 5.2.

2.4 An IDS checking program to support Supp. Comps.: Modification of Kawabata program or a New program ?

- IRG used to use Kawabata program to check IDS and enforce a 5% rule. With the introduction of Supp. Comps., IRG need to use an IDS check program to support Supp. Comps. in addition.

- I don't know if IRG plans to use a modified Kawabata program or to use a new program to support Supp. Comps. I can think of the following two broad options:

- Option 1: Modification of Kawabata IDS check program to support Supp. Comps.?

In this case, it need be discussed between IRG and Mr. Kawabata whether he will modify/improve his IDS check program to support Supp. Comps.

- Option 2: A new IDS check program to support Supp. Comps.?

Henry's online tool seems to support Supp. Comps. Currently, the tool seems to work in an interactive mode. I don't know if the tool also works in a batch mode which will probably be needed to check thousands of IDS lines in a batch mode.

A program other than Kawabata and Henry programs could be a candidate.

- It is suggested that IRG announce a feasible plan regarding this issue as a recommendation at this meeting.

- If IRG cannot secure an IDS check program (maybe working in a batch mode?) to support Supp. Comps., probably IRG cannot fulfill the objective of Supp. Comps. (i.e., better machine checking of IDS as mentioned in IRG N2225) and, therefore, it may not be desirable to allow Supp. Comps. to be used in IDS for WS2021 submission. If that is the case, Supp. Comps. in ws2017-ids.txt could be replaced with a full-width question mark U+FF1F ? so that the ids file can be used by an old IDS check program not supporting Supp. Comps.

2.5 Modification of IDS-relevant portions in ISO/IEC 10646

- I wonder if Supp. Comps. are intended to be used within IRG only. If that is the case,
a) it is suggested that such intent need be clearly mentioned in Supp. Comps. document and/or IRG PnP; and

b) the rest of this section might not seem relevant.

- The rest of this section assumes that Supp. Comps. are intended to be used not just within IRG but also outside of IRG.

- I wonder if IRG will propose to modify IDS-relevant clause (probably clause 1.2, Informative Annex I) in ISO/IEC 10646 to reflect that Supp. Comps. in IRG N2225 can be used in IDS (see Appendix 1 at the end of this document).

- Currently, the definition of IDS in the latest ISO/IEC 10646 does not seem to allow Supp. Comps (i.e., an HTML char. entity syntax) as its component.

Note. Private Use Character (PUC), or PUA (Private Use Area) character, can be a DC (Description Component) in IDS. However, probably Supp. Comps. cannot be considered as PUA char unless Supp. Comps. are assigned a UCS cp (see clauses 1.2 of Annex I and 7.3.5 in ISO/IEC 10646, 6th ed. which are shown in Appendices 1 and 2 at the end of this document). In other words, instead of using "&H7-01;" (a string composed of 7 chars), if we assign this Supp. Comp. to, for example, a PUA char U+0F1234 and use just one char (i.e., one UCS cp, U+0F1234) in IDS, it will be O.K. (Please correct me if I am wrong.)

2.6 Modification of some old IDS's for CJK main and Ext. A ~ G

- I wonder if, every time new Supp. Comps. are accepted by IRG, IRG will review and, if necessary, update some old IDS's for CJK main and Ext. A ~ G.

- Unless IRG updates some old IDS's whose IDS can be better described using new Supp. Comps., probably IRG cannot fulfill the objective of better machine checking of IDS (as mentioned in IRG N2225).

- In other words, IDS's of those chars whose IDS can be better described by utilizing new Supp. Comps. need be modified/updated (for example, by replacing a full-width question mark U+FF1F ? with a new Supp. Comp.) so that an IDS check program can do better IDS checking (of course, assuming that the program supports Supp. Comps.). According to ISO/IEC 10646, U+FF1F is used to represent an otherwise undescribed DC (Description Component) in IDS.

- If IRG wants to review old IDS's and, if necessary, modify some old IDS's, I am concerned that such work might become a burden to IRG which requires much time and effort.

2.7 Supp. Comps. in WS2017 and in WS2021

- If IRG wants to use Supp. Comps., it seems essential that

- a) IRG uses a program supporting Supp. Comps. and
- b) IRG reviews old IDS's of CJK main and Extensions and, if necessary, modifies/updates some old IDS's by using Supp. Comps. so that we can do better machine checking of IDS (please correct me if I am wrong).

- At the time of WS2017 submission, Supp. Comps. were allowed in IDS and Kawabata program was used. I don't know if Kawabata prog. supported Supp. Comps. at that time. This point should have been discussed when IRG allowed Supp. Comps. to be used for WS2017 submissions.

(Personally, I did not notice Supp. Comps. in WS2017 IDS until recently since KR did not use Supp. Comps. in its WS2017 submission.)

It is suggested that, if IRG is going to allow Supp. Comps. in WS2021 submissions, IRG adopt a recommendation RE: a feasible plan to use an IDS check prog. supporting Supp. Comps. and to review old IDS's and modify some old IDS's. (Please correct me, if I am wrong.)

2.8 An alternative approach

- I can think of an alternative approach which will require more time and effort but will cause less confusion. (I do not claim that this alternative is the best or will work without any problem.)

a) The set of Supp. Comps. is initially created by analyzing not only CJK main but also all CJK Extensions. By doing so, the set of Supp. Comps. will become less unstable and the need for NEW Supp. Comps. will be decreased.

b) The set of Supp. Comps. ver. 2020 that will be used for WS2021 submission is fixed before WS2021 submission.

In other words, new Supp. Comps. not in the set of Supp. Comps. ver. 2020 are not allowed in the WS2021 submissions.

c) Once a new version of the set of Supp. Comps. is fixed, the old IDS's are reviewed and some old IDS are modified accordingly so that we can fulfill better machine checking of IDS.

A table comparing the current and the alternative approach is shown below:

currently: an incremental approach	an alternative approach
the current set of Supp. Comps. are mostly from CJK main (URO) -> The set is unstable e.g., size of the set: 45 -> 122 -> 180 (?)	the set of Supp. Comps. are extracted from CJK main and all Extensions -> The set is not so unstable (disadvantage: It takes more time.) e.g., size of the set: 500 (assumption) -> 530 -> 550 ...
NEW Supp. Comps. allowed in WS2021 submissions	the set of Supp. Comps. ver. 2020 is fixed and used for WS2021 submission; (i.e., new Supp. Comps. not in ver. 2020 are NOT allowed in WS2021 submissions)
?	some old IDS's are modified based on Supp. Comps. ver. 2020
submitters can suggest and use new Supp. Comps. in WS2025 submission (i.e., no restriction)	the set of Supp. Comps. ver. 2024 is fixed (possibly by adding new Supp. Comps.) and used for WS2025 submission (i.e., new Supp. Comps. not in ver. 2024 are NOT allowed in WS2025 submissions)
?	some old IDS's are modified based on Supp. Comps. ver. 2024

Note. In Appendix 3, another issue is discussed and a suggestion is made. That issue does not seem directly related with Supp. Comps. issue, but is related with IDS issue in general, which explains why that issue is discussed in Appendix 3.

* * *

Appendix 1. A part of Informative Annex I, ISO/IEC 10646

ISO/IEC 10646:2020 (E)

Annex I (informative) Ideographic description characters

I.1 General

An Ideographic Description Character (IDC) is a graphic character, which is used with a sequence of other graphic characters to form an Ideographic Description Sequence (IDS). Such a sequence may be used to describe an ideographic character which is not specified within this document.

The IDS describes the ideograph in the abstract form. It is not interpreted as a composed character and does not imply any specific form of rendering.

NOTE – An IDS is not a character and therefore is not a member of the repertoire of this document.

I.2 Syntax of an ideographic description sequence

An IDS consist of an IDC followed by a fixed number of Description Component (DC) organized in subgroups corresponding to script category, such as CJK ideographs or Tangut ideographs. An IDS should only use items belonging to a single subgroup, to clarify the script that it belongs to. The subgroups and their contents are as follows:

- CJK IDS subgroup, including DCs which may be of any one of the following:
 - a coded CJK ideograph, which consists of any coded character from the CJK UNIFIED IDEOGRAPHS blocks or the CJK COMPATIBILITY IDEOGRAPHS blocks,
 - a coded CJK radical, which consists of any coded character from the CJK RADICALS block or the KANGXI RADICALS block,
 - a coded CJK stroke, which consists of any coded character from the CJK STROKES block,
 - the character FF1F FULLWIDTH QUESTION MARK to represent an otherwise undescribed DC,
 - a private use character (as long as the interchanging parties have agreed that the particular private use character represents a particular CJK ideograph or component of a CJK ideograph),
 - another CJK IDS.

Appendix 2. Clause 7.3.5 of ISO/IEC 10646

7.3.5 Private use characters

Code points from E000 to F8FF in the BMP are reserved for private use. All code points of Plane 0F and Plane 10 – except for noncharacter code points FFFFE, FFFFF, 10FFFE, and 10FFFF – are reserved for private use.

Private use characters are not constrained in any way by this document. Private use characters can be used to provide user-defined characters. For example, this is a common requirement for users of ideographic scripts.

NOTE – For meaningful interchange of private use characters, an agreement, independent of this document, is necessary between sender and recipient.

Appendix 3. Improving information about "Equivalence Database" (IDS) at IRG meeting web sites

- There are two links for IDS data (or "Equivalence Database") at IRG #56 web site as of 2021-02-28 as shown below

(<https://appsrv.cse.cuhk.edu.hk/~irg/irg/irg56/IRG56.htm>)

Equivalence Database(Updated on 2014.08.20): <https://github.com/cjkvi/cjkvi-ids>
../irg31/IRGN1277 Appendix: http://www.itscj.ipsj.or.jp/domestic/sc02/irg-files/CJK_D_attributes/IRG1277_attachment.zip

. the first link is

<https://github.com/cjkvi/cjkvi-ids>

. the second link is

../irg31/IRGN1277 Appendix:

http://www.itscj.ipsj.or.jp/domestic/sc02/irg-files/CJK_D_attributes/IRG1277_attachment.zip

- The second link seems broken. Based on the file name, the file seems to contain IDS for CJK Ext. D. Since ids.txt on the page pointed to by the first link already contains IDS for CJK Ext. D, the second link seems unnecessary and, therefore, this link can safely be deleted.

- When I follow the first link, I can see about 9 .txt files which seem to have some kind of IDS data. It is suggested that, at IRG meeting #56 web site, the file names that will be used for IDS checking by Kawabata prog. for WS2021 be listed.

- It seems that the following two files might be used for checking IDS of WS2021:

. ids.txt (IDS's for CJK Unified main and Ext. A ~ F)

. ws2015-ids.txt (IDS's for WS2015; submitter's source reference used instead of UCS cp)

- It is suggested that the date in "Equivalence Database(Updated on 2014.08.20.)" be modified to "2017" based on the first line of the second link:

Copyright (c) 2014-2017 CJKVI Database

- As a summary, it is suggested that the current info about IDS (Equivalence Database) at IRG meeting web site be modified as shown below (or something like this):

start of current text

Equivalence Database(Updated on 2014.08.20): <https://github.com/cjkvi/cjkvi-ids>

../irg31/IRGN1277 Appendix:

http://www.itscj.ipsj.or.jp/domestic/sc02/irg-files/CJK_D_attributes/IRG1277_attachment.zip

end of current text

==>

start of suggested text (text only in bold will be shown to users)

Equivalence Database: <https://github.com/cjkvi/cjkvi-ids>

*1) **ids.txt (CJK main and Ext. A ~ F; last updated 2017): 88397 lines excluding the first two comment lines.***

*2) **ws2015-ids.txt (last updated 2016): 5065 lines excluding the first one comment line (submitter's source reference used instead of UCS cp)***

*3) **ws2017-ids.txt** (if and when provided)*

end of suggested text (text only in bold will be shown to users)

- If the suggested text is too long, it could be shortened.

- Note. This issue does not seem directly related with Supp. Comps. issue, but is related with IDS issue in general, which explains why this issue is discussed in Appendix 3.

* * *

Subject: A few considerations on CJK Supplementary Components for IDS (Part 2)

Date: 2021.03.12.

Author: KIM, Kyongsok (Republic of Korea)

Status: Individual Contribution

After submitting IRG N2464 dated on 2021.03.08. to IRG, I dug old IRG N documents related with Components used in IDS and found new facts, which explains why I write part 2 of IRG N2464 (the document number of this document is IRG N2464_2).

The last Subsection number in Section 2 in IRG N2464 is 2.8.

This document starts with Subsection number 2.9 following 2.8 in IRG N2464.

2.9 About 700 CDP glyphs (components)

- In IRG N1939, "Regarding to CDP characters used in IDS database", KAWABATA Taichi, 2013-05-17, CDP components are introduced in response to IRG Resolution M39.9. Some parts of the document are quoted below:

*This document is a **response to IRG M39.9 2**. This document describes the **CDP glyphs** used in IDS database [3] as an "alternate" for **unencoded component**, and propose several ways to handle unencoded components in the IDS.*

CDP glyphs in IDS data

*Attached glyph list shows the list of glyphs used in the IDS data [2] that are used in IDS data [3]. These glyphs are taken from **EUDC (End User Defined Characters) font** provided from **Chinese Document Processing (CDP) Laboratory of Academia Sinica** (<http://cdp.sinica.edu.tw/>). They are denoted as `&CDP-XXXX;` (an entity reference style notation) in IDS data 3, that is a practice of CHISE Project (<http://www.chise.org/>), where original IDS data is created and distributed.*

A list of about 700 CDP components is attached to that document. Some CDP components are shown below:

UCS	Big5	char	IDS	num	subtraction
F137	854B	𠄎	𠄎 𠄎 从	1	=两-一
F138	854C	𠄎	𠄎 𠄎 土	0	
F139	854D	𠄎	𠄎 𠄎 效	2	=兩-工
F13A	854E	𠄎	𠄎	1	=兩-門-吞
F13D	8551	𠄎	𠄎 𠄎 木 𠄎 𠄎 八	1	=劃-田-一-一 𠄎

I am not an Hanja expert to evaluate CDP components. I just wonder if CDP could be a good starting basis for Supp. Comps. I don't know the IRG conclusion regarding whether to allow CDP chars in IDS.

2.10 A list of PUA chars (UCS cp. and corresponding glyph) ?

In WG2 N4241, "Information in support of N4234 (L2/12-087) to demonstrate extensive use of PUA in common IDS data" by Dr. Ken Lunde, 2012-02-14, we can see the following statement:

The list below indicates how many characters in each CJK Unified Ideograph block use private use characters in their IDSes, based on the current version of the IDS database:

URO: 480
Extension A: 107
Extension B: 1,972
Extension C: 55
Extension D: 14

If IRG allowed MB to use some designated PUA chars in IDS within IRG, the list of these PUA chars need be maintained somewhere. I don't know if there is a set of such PUA chars.

2.11 Circled digits in IDS ?

Circled digits are found in ids.txt at <https://github.com/cjkvi/cjkvi-ids>, as shown below:

U+5DE4 巖 𠄎𠄎𠄎𠄎𠄎𠄎 [GJK] 𠄎𠄎𠄎𠄎 [T]
U+689F 梟 𠄎 ① 木
U+83EF 華 𠄎 ++ ①
U+88CA 裊 𠄎 ① 衣
U+9115 鄉 𠄎 乡 𠄎 𠄎 [GTKV] 𠄎 乡 ① 𠄎 [J]
U+9B1B 鬚 𠄎 𠄎 𠄎 𠄎 ①

Actually Mr. Kawabata mentioned the possibility of using circled digits, ① ~ ⑩, in IDS in his document IRG N1939 (see 2.9 above for details).

I wonder if IRG allowed/allows MB to use circled digits in IDS (within IRG only?).

2.12 A full-width question mark, PUA chars, circled digits, CDP components, Supp. Comps., etc.

As we saw above, at least five possibilities were proposed to represent unencoded components: 1) a full-width question mark, 2) PUA chars, 3) CDP Components, 4) circled digits, 5) Supp. Comps..

It is suggested that IRG discuss whether to allow to use only one method OR two or more methods in MB submissions and in old IDSes.

A full-width question mark and PUA chars conform to Informative Annex I whereas CDP components, circled digits, and Supp. Comps. do not.

2.13 More than one IDS for one char.

- It is suggested that IRG discuss whether to allow more than one IDS for one CJK char and, if the answer is yes, specify a syntax to express two or more IDSes for one char.

* * *

Source:	BAI Yi
Title:	Feedback on A few considerations on CJK Supplementary Components for IDS
Status:	Individual Contribution on IRG #56
Action:	To be considered by IRG

1. In Section 2.4, it is proposed that either Kawabata-san's program or a new program can be used in IDS checking. Currently I'm developing a new IDS check system that can fulfill requirements listed in Annex B of PnP, additionally,
1) it can do "fuzzy search" by applying UCV rules to IDS to find potential unifiable characters. We've found unifiable characters in the working set 2017 by this tool.

Unification

Sn	Image/Source	Comment Type	Description
00145	<p>00145 魚</p> <p>人 9.11 13 □ (3) V-F0B0C</p> <p>Trad □イ魚</p>	Unification	<p>29D51 僂 僂 僂</p> <p>魚 195.2 UCS2003 GKX-1465.12 T5-444F</p> <p>Unifiable with 僂 under UCV rule 322.</p> <p>322 魚 臭 臭</p>

Fig 1. Potential unifiable character found in ws2017 by applying UCV rules
<https://hc.jsecs.org/irg/ws2017/app/index.php?id=00145>
Also, about 200 possible unifiable characters are found in existing character sets.
<https://github.com/eisoch/irg/issues/28#issuecomment-781768239>

2) it can be extended to include Supplementary Components in checking.
I think, if possible, this system can work along with Kawabata-san's program in future working sets to provide better machine checking on IDS and additional advice to reviewers.

2. In Section 2.6, it is proposed that when Supplementary Components are updated, old IDSes for existing characters should also be updated. In addition to that, when a new extension set is published, of course, characters in that extension set can also be used in

IDSes. For example, 市 (U+30009), which was encoded in extension G, has identical

shape of &H5-01. Thus all IDS including this component can be updated. After an update on the working set, or after a new extension set is released, working in batch mode, our IDS system can check the IDS of all encoded characters and characters working sets, to provide advice about optimizing the IDS.

03966	<p>03966 褱</p> <p>衣 145.12 18 □ (1) UTC-03143</p> <p>Trad 目&H5-01;一石衣</p>	IDS	As 市 has been encoded the IDS can be optimized to 目市一石衣.
-----------------------	--	-----	--

Fig 2. Advice on optimizing IDS after checking

(End of document).

2021/03/13

Feedback to IRG N2464

suzuki toshiya

Abstract

Basically, I support the points raised by Professor Kim, there are some points to be documented in IRG PnP and Working Document Series. I try to answer a few questions and write a few comments (mainly for the part 1, and partially for the part 2 of IRG N2464).

Comments on Kawabata-san's program

The questions I can try to answer are not so many, so please let me put the number for this document.

Q1: The equivalence / unification candidate program is ready to handle Supp. Comps.?

A1: Yes, but depends on the quality of the maintenance condition of the past ids.txt. Please think about the situation:

IRG MB A submits a character A-0123, whose IDS is 𠄎示&H7-01;

IRG MB B submits a character B-2345, whose IDS is 𠄎禛 &H7-01;

The program by Kawabata-san would find the similarity between A-0123 and B-4567, with no knowledge about the glyph shape of &H7-01;

But, another situation, like,

IRG MB X submits a character X-6789, whose IDS is 𠄎目&H3-04;

The program by Kawabata-san would be hard to find its similarity with 具, because the program and its currently predefined IDS data have no knowledge about the glyph shape of &H3-04;. To improve this situation, we should update the IDS for 具 - currently it has 2 IDSs (𠄎&CDP-8BA5; 八 and 𠄎目一八).

Yet I'm unsure whether addition of the 3rd IDS (𠄎目&H3-04;) is enough, or we should replace old 2 IDSs by new one.

This question is strongly related with the question raised in 2.6 "IDS for URO, ExtA ~ ExtG should be updated?". I think, if we include some entries like a macro symbol definition (of programming language C), like,

&H3-04; 𠄎一八

might improve the situation.

Q2: IRG will propose to modify IDS-relevant clauses in ISO/IEC 10646 to reflect Supp. Comps. in IRG2225?

A2: Maybe it would be dealt with as out of scope of ISO/IEC 10646. ISO/IEC 10646 or Unicode can discuss the plain text without markup. But the syntax of Supp. Comps. is clearly XML-like. Even if we try to write such in ISO/IEC 10646, we would be asked a question “Are you saying this is plaintext? It looks like a portion of the marked-up language, the syntax of the marked-up language does not fit to ISO/IEC 10646”. The better way is trying to find an appropriate organization to define the standard for the marked-up language for the light-weight ideographic structure description, if IRG has to take the responsibility about the usage of IDS data out of IRG.

Comments on other issues

- I agree with the requirement of the update of IRG PnP (or IRG Working Document) to clarify the syntax of IDS which can be used in the submission to IRG.
- I agree with the requirement of the form for IRG MBs to propose/submit new Supp. Comps, to minimize the cost of the consolidation and review works.
- I agree with the requirement of the definition of the rule to name new Supp. Comps used in the working set submissions, to minimize the cost of the review works.
- About a few “new” syntaxes proposed by Kawabata-san (IRG N2464 part 2, 2.11 and 2.12), IRG is expected to issue an official recommendation for the future submissions, which format is the best for IRG. It would prevent the case that the different syntaxes are used in multiple submissions, and it would be useful to reduce the cost of the consolidation & review.
- The feedback from Bai Yi demonstrates the automation of the update of the IDS data after the standardization of new extensions in CJK Unified Ideograph. I wish if it can be used to estimate “how many IDSs could be optimized by the introduction of this new Supp Comp”.

(end of document)