

## Korea JTC1/SC2, Committee on Character Codes

---

Author: KIM, Kyongsok

Date: 2012.11.13.

Status: Expert's contribution

Subject: Clarifying First Stroke (FS) in IRG PnP

### 1. Background

– Since there was confusion as to the correct definition of First Stroke and there were inconsistencies in IRG PnP RE: First Stroke, several e-mail messages were exchanged RE: First Stroke in IRG PnP. (see the attachments)

– Furthermore, suddenly MBs are requested to provide FS info in ExtF submission form, it became important to clearly specify FS in IRG PnP to avoid possible confusion.

### 2. Suggestion

#### 2.1 The author supports Mr. Chen's suggestion as shown below:

*from: chen-zhuang chenzh-zhuang@163.com*

*date: Thu, Jun 28, 2012 at 7:13 PM*

*subject: Re:Re[2]: Re: [irgeditors] Question Set #2 RE: FS (First Stroke) value*

*... we agreed to let the Technical Editor to choose FS for every character.*

*BTW, I would like to support the Chinese version of FS because this order was recognized by the Ministry of Education of PRC and has been implemented in the mainland of China for many years.*

2.2 Therefore, the author suggests that China (Mr. Chen and/or Ms. Wang) modify IRG N1823 Draft3 IRG PnP appropriately while clearly explaining some issues raised below:

1) issue 1: whether FS is the first stroke of non-radical component (vs. the first stroke of the total character)

2) issue 2: the Dic Version A is the CORRECT FS values.

(Currently First strokes specified in Annex K of IRG PnP (IRG N1823Draft3) corresponds to UCS version, which is INCORRECT.)

UCS version		Dic Version A
<b>First stroke</b>		<b>First stroke</b>
— 1 One		— 1 横 heng
2 Line		2 豎 shu
\ 3 Dot		丿 3 撇 pie
丿 4 Slash		\ 4 点 dian
乙 5 Second		㇏ 5 折 zhe

3) issue 3: consistent names of First Strokes both in English and in Hanzi.

4) issue 4: whether the direction and the length are important on deciding if FS value is 3 (slash) or 4 (dot).

5) issue 5: which of the following four documents are valid and therefore can be referenced in IRG PnP:

IRG N881 (I am not sure if it is still valid...)

IRG N954

IRG N954AR

IRG N1105

\* \* \*

## Attachments (Hangeul text can be safely ignored...)

### Attachment 1

from: KIM Kyongsok gimgs0@gmail.com  
to: irgeditors@ml.comp.polyu.edu.hk  
cc: KIM Kyongsok <gimgs0@gmail.com>  
date: Wed, Jun 27, 2012 at 6:07 PM  
subject: A question RE: FS (First Stroke) value

Dear editors,

1. I have difficulty understanding FS (First Stroke) values. Specifically, I cannot figure out why FS value of 3, 4, or 5 is assigned to the examples shown below.

#### First stroke

- 一 1 One
- 丨 2 Line
- 丶 3 Dot
- 丿 4 Slash
- 乙 5 Second

<div>㐓</div> <div>02208</div> <div>11</div> <div>3</div>	<div>㐓</div> <div>02209</div> <div>11</div> <div>3</div>	<div>㐓</div> <div>02210</div> <div>11</div> <div>3</div>
<div>㐓</div> <div>02211</div> <div>11</div> <div>3</div>	<div>㐓</div> <div>02213</div> <div>11</div> <div>4</div>	<div>㐓</div> <div>02214</div> <div>11</div> <div>4</div>
<div>㐓</div> <div>02215</div> <div>11</div> <div>5</div>	<div>㐓</div> <div>02236</div> <div>13</div> <div>5</div>	<div>㐓</div> <div>02224</div> <div>12</div> <div>4</div>
<div>㐓</div> <div>02211</div> <div>11</div> <div>3</div>	<div>㐓</div> <div>02214</div> <div>11</div> <div>4</div>	<div>㐓</div> <div>02224</div> <div>12</div> <div>4</div>
<div>㐓</div> <div>02215</div> <div>11</div> <div>5</div>	<div>㐓</div> <div>02236</div> <div>13</div> <div>5</div>	<div>㐓</div> <div>02237</div> <div>13</div> <div>5</div>

2.

2.1 For 02208, 02209 and 02210, FS value 4 (Slash) seems more appropriate.  
2.2 For 02211, I wonder if we ignore the direction of a dot on deciding FS value.

- In the leftmost column, the dot is drawn from NW(northwest) to SE (southeast).

However, the dot of 02211 is drawn from NE to SW.

3.

3.1 For 02213, FS value 3 (Dot) seems more appropriate if we ignore the direction of a dot.

3.2 For 02214 and 02224, FS value 3 (Dot) seems more appropriate.

4.

4.1 What does FS value "5 Second" mean?

4.2 I wonder why FS value 5 (Second) was assigned to 02215, 02236 and 02237?

Thanks in advance. Please help me...

I am so much confused...

KIM, K.

\* \* \*

참고. UCS에 U2000 - U2004에 나오는 강시 부수 첫 다섯 개

### **Kangxi radicals**

2F00	一	KANGXI RADICAL ONE $\approx$ 4E00 一
2F01		KANGXI RADICAL LINE $\approx$ 4E28
2F02	丶	KANGXI RADICAL DOT $\approx$ 4E36 丶
2F03	丿	KANGXI RADICAL SLASH $\approx$ 4E3F 丿
2F04	乙	KANGXI RADICAL SECOND $\approx$ 4E59 乙

Attachment 2

from: xiaomwang2006 xiaomwang2006@163.com  
to: KIM Kyongsok <gimngs0@gmail.com>  
date: Wed, Jun 27, 2012 at 7:41 PM  
subject: Re:Fwd: A question RE: FS (First Stroke) value

Dear KIM:

Thanks, I have returned safe to my home.

The FS is not the first stroke of the total character, but the non-radical component. I don't know if I answer your question?

Best Regards

Wang Xiaoming

### Attachment 3

- 따라서 IRG P&P에 나와 있는 3 번이 점(dot), 4 번이 세로줄(slash)은 잘못됐음

from: chen-zhuang chenzh-zhuang@163.com  
to: KIM Kyongsok <gimgs0@gmail.com>  
cc: irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
date: Wed, Jun 27, 2012 at 10:33 PM  
subject: Re:[irgeditors] A question RE: FS (First Stroke) value



Prof. Kim,



The First Stoke Value is as follows:



1. horisonal line
2. vertical line
3. slash
4. dot
5. turning stoke

You mixed 3 and 4.

For the 3 characters with FS5, I marked in the attachment.

	0291.181	
		
	11	5
02215		

	0292.231	
		
	13	5
02236		

	0292.231	
		
	13	5
02237		

#### Attachment 4

from: KIM Kyongsok gimgs0@gmail.com  
to: chen-zhuang <chenzh-zhuang@163.com>  
cc: irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
date: Wed, Jun 27, 2012 at 11:21 PM  
subject: [irgeditors] Question Set #2 RE: FS (First Stroke) value


Dear Mr. Chen Zhuang:

Thanks very much for your helpful comments.  
Now I came to understand FS in many of the examples I mentioned.  
However, I still have a few questions.


\* For details and Hanzi chars, plz see the attached PDF file.

<b>First stroke</b>			
→	1 One		
	2 Line		
↘	3 Slash		
·	4 Dot		
乙	5 turning stroke		

	0291.181
	↘
	11   3
02211	

	0291.181
	↘
	11   4
02213	

4. Do we ignore the direction of a slash/dot on deciding whether FS value is 3 (slash) or 4 (dot).

- In other words, which of the direction and the length is important on deciding whether FS value is 3 (slash) or 4 (dot).

- The answer to #4 may answer #5 and #6 below.

Thanks in advance.

KIM, K.

\* \* \*

Attachment 5

from: chen-zhuang chenzh-zhuang@163.com  
to: KIM Kyongsok <gimsgs0@gmail.com>  
cc: irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
date: Thu, Jun 28, 2012 at 8:41 AM  
subject: Re:[irgeditors] Question Set #2 RE: FS (First Stroke) value

Prof. Kim,

Yes, we ignore the direction of a slash/dot on deciding whether FS value is 3 (slash) or 4 (dot).

For 02211, the FS is selected 3 (slash) because it is a part of “白”, then the next FS is selected 4 (dot) because it is a part of “必”.

Chen Zhuang



Attachment 6

from: rkfyan@ogcio.gov.hk  
to: chen-zhuang <chenzh-zhuang@163.com>, KIM Kyongsok  
<gimsgs0@gmail.com>  
cc: irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
date: Thu, Jun 28, 2012 at 12:40 PM  
subject: Re:[irgeditors] Question Set #2 RE: FS (First Stroke) value

Dear Chen Zhuang, Prof. KIM & editors,

To follow on your email exchange, we want to point out that a new annex – Annex K: A List of FS (First Strokes) has been added in the IRGN1823Draft2 – IRGPnP Update. However, stroke 3 & 4 shown in the annex do not tally with Chen Zhuang's note in the earlier email. I believed the figure was 'extracted' somewhere. It seems that in some previous document, stroke 3 & 4 may be recorded differently.

Regards,  
Retarkgo

Attachment 7

from: Bear Tseng beartsn@gmail.com  
to: rkfyan@ogcio.gov.hk  
cc: chen-zhuang <chenzh-zhuang@163.com>, KIM Kyongsok  
<gimngs0@gmail.com>,  
irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
date: Thu, Jun 28, 2012 at 12:52 PM  
subject: Re: [irgeditors] Question Set #2 RE: FS (First Stroke) value

Dear Retarkgo,

According to the sequence of the first 5 Kangxi Radicals, i.e., U+02F00 to U+02F04, the dot shall precede than slash. You are correct.

Best regards,

Attachment 8

from: rkfyan@ogcio.gov.hk  
to: Bear Tseng <beartsn@gmail.com>  
cc: chen-zhuang <chenzh-zhuang@163.com>, KIM Kyongsok  
<gimgs0@gmail.com>,  
irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
date: Thu, Jun 28, 2012 at 1:00 PM  
subject: Re: [irgeditors] Question Set #2 RE: FS (First Stroke) value

Dear Bear,

I'm not sure which version is correct and just want to point out our observation. I think a consistent stroke reference should be used and clearly stated in the PnP for reference.

Regards,  
Retarkgo

Attachment 9

from: Bear Tseng beartsn@gmail.com  
to: rkfyan@ogcio.gov.hk  
cc: chen-zhuang <chenzh-zhuang@163.com>, KIM Kyongsok  
<gimsgs0@gmail.com>,  
irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
date: Thu, Jun 28, 2012 at 1:10 PM  
subject: Re: [irgeditors] Question Set #2 RE: FS (First Stroke) value

Dear Retarkgo,

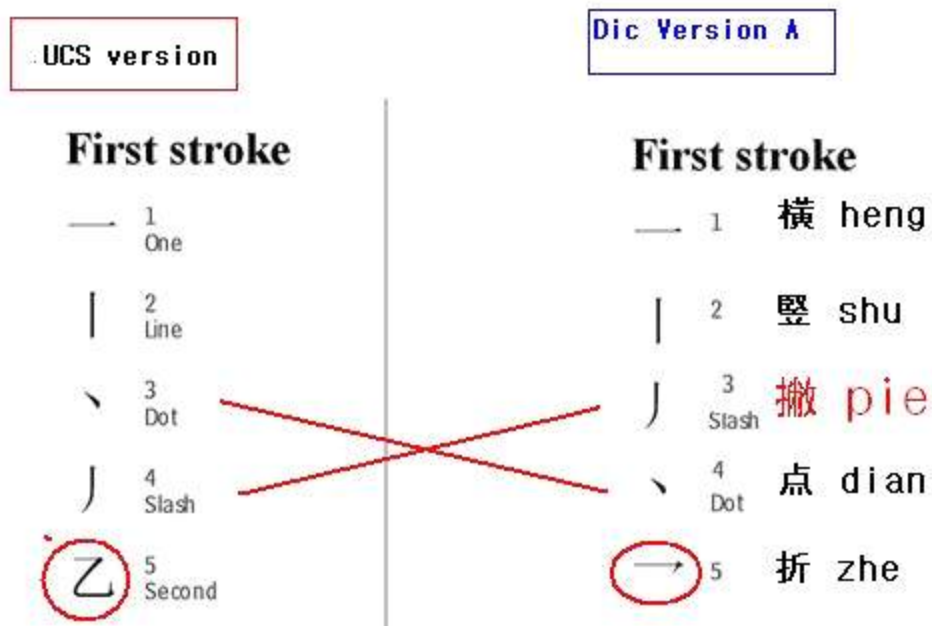
I said that because the 5 strokes were actually derived from the first 5 Kangxi Radicals by earlier Chinese librarians in Peking Library.

Best regards,

Attachment 10

from: KIM Kyongsok gimgs0@gmail.com  
to: irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
cc: KIM Kyongsok <gimgs0@gmail.com>  
date: Thu, Jun 28, 2012 at 2:50 PM  
subject: Re: [irgeditors] Question Set #2 RE: FS (First Stroke) value  
Dear editors:

1. The order and name of five strokes in FS chart in IRG PnP (a left chart below) seems to come from UCS. I will call it "UCS version".



2. I got the right (NOT CORRECT) chart from a dictionary that I have.  
- I will call it "Dic. version A".

- Other dictionaries may have different order/shape/name.  
Comments are welcome  
(esp. from people in China. Plz check your dictionaries).

- I noted three things:

- 1) The order of dot and slash are different (reversed).
- 2) The fifth shape (stroke) is different.
- 3) The names are different.

3. IRG needs to discuss and decide on ONE agreed-on FS chart so that

we can use order/shape/name consistently.

- We also need to check whether the order was consistent in several versions.

- For example, between main/A/B/C/D/E;  
and between different versions of the same block (say, ExtE v6.0, v7.0, etc)

Thanks.

KIM, K.

---

Attachment 11

from: xiaomwang2006 xiaomwang2006@163.com  
to: KIM Kyongsok <gimsgs0@gmail.com>  
date: Thu, Jun 28, 2012 at 3:19 PM  
subject: Re:Re: [irgeditors] Question Set #2 RE: FS (First Stroke)  
value

Dear editors.

Mr.Chen Zhuang is right.The First Stoke Value is as follows:

1. horizontal line "—"
2. vertical line"|"
3. slash"丿"
4. dot"丶"
5. turing stoke"乙"

We have been so used in CJK E always. Considering the continuity, I hope that we should not change it.

Best regards  
Wang xiaoming

Attachment 12

from: KIM Kyongsok gimgs0@gmail.com  
to: irgeditors@ml.comp.polyu.edu.hk  
cc: KIM Kyongsok <gimgs0@gmail.com>  
date: Thu, Jun 28, 2012 at 4:44 PM  
subject: Re: Re: [irgeditors] Question Set #2 RE: FS (First Stroke)  
value

Dear editors:

1. I guess it may be good to merge the following four docs into one:

IRG N881 (I am not sure if it is still valid...)  
IRG N954  
IRG N954AR  
IRG N1105

2. There seems some inconsistency among these four docs.

As an example, according to N881, the direction of slash and dot IS IMPORTANT whereas Mr. Chen said the direction is IGNORED.

I am not sure if N881 is still valid for our current work.

Another example. Glyph2 in N881 seem confusing since they may be different from what we adopt these days.

Thanks.

KIM, K.

---



Attachment 12

from: satoshi.yamamoto.yd@hitachi.com via ml.comp.polyu.edu.hk  
to: gimgs0@gmail.com  
cc: irgeditors@ml.comp.polyu.edu.hk  
date: Thu, Jun 28, 2012 at 5:32 PM  
subject: Re[2]: Re: [irgeditors] Question Set #2 RE: FS (First Stroke) value

Dear Prof. Kim,

Actually I don't know what was discussed on IRG N881 or N954 because I was not attending IRG at that time.

As I said this morning, IRG N881 is the first explanation on FS. However, at the practical view point, I guess IRG found so many variations in the actual glyph shape after collected characters. In other words, there are various directions for the specific stroke in the same components depend on regions. So IRG needed to normalize FS or SC for some popular components. As a result, IRG N954AR was published at IRG #22 (amended by IRG N1105).

So I understand we need to refer IRG N881 N954AR as a principle because IRG N954AR is not exhaustive. We don't need writing about the definition of FS strictly because we have agreed at the last meeting to ask technical editor to produce and maintain this information.

What do you think ?

—

YAMAMOTO Satoshi <mailto:satoshi.yamamoto.yd@hitachi.com>  
Software Engineering Dept., Hitachi, Ltd.

\* Sorry, typo.

So I understand we need to refer "IRG N881" as a principle because IRG N954AR is not ...

Attachment 13

from: KIM Kyongsok gimgs0@gmail.com  
to: satoshi.yamamoto.yd@hitachi.com  
cc: irgeditors@ml.comp.polyu.edu.hk  
date: Thu, Jun 28, 2012 at 5:46 PM  
subject: Re: Re[2]: Re: [irgeditors] Question Set #2 RE: FS (First Stroke) value

Mr. YAMAMOTO wrote:

> We don't need writing about the definition of FS strictly because we have  
> agreed at the last meeting to ask technical editor to producu and  
maintain  
> this information.  
>  
> What do you think ?  
>  
> --  
> YAMAMOTO Satoshi <mailto:satoshi.yamamoto.yd@hitachi.com>

Dear Mr. Yamamoto:

Right, we agreed that tech. editor will assign FS value.  
Other people need not worry about assigning FS value. ^ - ^

I agree with Mr. Yamamoto that we may not need a PERFECT rule on FS.  
However, there seems some INCONSISTENCY among the four docs.

Therefore, we may need a CONSISTENT (may not be PERFECT though) rule  
so that tech. editors and IRG members understand how FS is assigned.

Just my thought.

KIM, K.

---

from: KIM Kyongsok gimgs0@gmail.com  
to: satoshi.yamamoto.yd@hitachi.com  
cc: irgeditors@ml.comp.polyu.edu.hk  
date: Thu, Jun 28, 2012 at 5:54 PM

One more thing:

My estimate is that merging four docs and removing a few inconsistencies

won't be too hard.

We need to agree how to remove a few inconsistencies (it won't be hard, I guess) and then need to edit docs.

Thanks.

KIM, K.

Attachment 14

from: chen-zhuang chenzh-zhuang@163.com  
to: "satoshi.yamamoto.yd" <satoshi.yamamoto.yd@hitachi.com>  
cc: gings0 <gings0@gmail.com>,  
irgeditors <irgeditors@ml.comp.polyu.edu.hk>  
date: Thu, Jun 28, 2012 at 7:13 PM  
subject: Re:Re[2]: Re: [irgeditors] Question Set #2 RE: FS (First  
Stroke) value

I agree to Yamamoto san too. It's unnecessary to focus on FS so far since we agreed to let the Technical Editor to choose FS for every character.

BTW, I would like to support the Chinese version of FS because this order was recognized by the Ministry of Education of PRC and has been implemented in the mainland of China for many years.

Chen Zhuang

\* \* \*