Date: Jan 26 2005

Request to add clarification concerning some HKSCS CJK Ideographs in Annex P of ISO/IEC 10646:2003

From: KAWABATA, Taichi, Lu Qin (experts contribution)

We would like to request to add the following paragraph in Annex P. of ISO/IEC 10646:2003.

9FB9, 9FBA, and 9FBB:
These three characters are intended to represent a component at a specific position of a full ideograph. To represent the ideograph of the same structure without a specific positional preference, 20509, 2099D and 0470C should be used respectively.

We request it because these pairs share the unifiable shape, and they may confuse people as to which code points they should use to encode their documents.

Background:
WG2 N2807 suggests that U+9FB9, U+9FBA, and U+9FBB share the unifiable shapes with U+20509, U+2099D and U+0470C, but they should be separated because the former are high or left half 'component', as shown in Figure 1.

However, the IRG has not yet reached the consensus whether the glyph placed at the specific half side should be treated differently from the one placed in the middle. Usually, in such a case, the characters are placed at the compatibility area, as U+FA5E and U+8279.

Additional Information:

For U+20089 and U+215D7 mentioned in N2807, Japan's JIS X 0213 has the characters 2-1-1 and 2-5-30, whose shapes are placed in the middle of the grid, as shown in the following figures (Screenshots are taken with adjacent character for metric comparison).
2-1-1 and 2-5-30 of JIS X 0213 correspond to U+20089 and U+215D7 of UCS, which places them at the "higher half", as shown in the following figures (screenshots are taken with adjacent characters for metric comparison).

This shows that the placement of the shape has not been an issue for the character representation. By accepting U+9FB9, 9FBA, and 9FBB it would mark for the first time that the placement of the shape may affect the choice of the code point. Thus we should put the above statement for the clarification.

Note:
JIS X 0213 introduces the following characters (2-3-7, 2-3-51, 2-88-75), as shown in the following figures.

And there are various fonts which implements JIS X 0213 shapes for these code points.