

Title: Request to Encode Variant Characters of Siddham Script

Source: Taichi Kawabata, Kiyonori Nagasaki, and Toshiya Suzuki

Status: Expert Contribution

Date: 21 Nov 2013

First of all, we appreciate the efforts of Unicode people on encoding Siddham script, and we're sorry that we submitted the comments on Siddham script on late time. Normally, appropriate Japanese people would review proposals earlier. However, when Kiyonori Nagasaki who got involved in the Japanese SC 2 committee as an expert of encoding of Buddhist CJK ideographs saw the proposal for the first time, it was too late to discuss issues about encoding. Although we discussed with experts of Siddham script, we didn't have enough time to solve potential problems in the proposal. So, we apologize for discussing intrinsic problems at this late stage.

We believe it is very important that Unicode includes some specific variant characters in Siddham script. As Siddham script has been handed down from master to pupil, there are several disunities on how Japanese people treats Siddham variants, and it is difficult to confirm all rules in Japan. One character in Siddham represents not only a phonetic value, but also a kind of bodhisattvas. Then, Japanese Siddham users would usually treat variant characters differently in order to distinguish bodhisattvas which have the same pronunciation. Naturally, they would expect that Siddham script should include these variant characters. If Unicode do not include such variant characters, they will seek other (probably non-Unicode) methods to encode and interchange Siddham characters. If such a thing happens, Japanese IT vendors may also adopt such a nonstandard encodings to satisfy their customers. We're very afraid that if Unicode were to ignore the variant characters in Siddham script, Japanese people would consider that Unicode is useless for such cultural materials, and were likely to address the problem by embracing their own new character encoding. However, such an approach would not sustain for the future, and it is obviously better to cooperate internationally. So, to avoid such situations, we believe it is important to be able to convey to them that Unicode is mindful of Japanese traditions of Siddham script, that is, regional diversity.

We strongly believe that the variant differences should be distinguished at the encoding level if such characters convey fundamental meaning and can be exchanged properly with others without font specification. As far as we have investigated, such cases occur

only in the small number of characters that we proposed in N4407R. If other cases would be found, we would be happy to add them in a revised or future proposal. Overall we would cordially request again that Unicode respects the traditional usage of Siddham script in Japan.

We would also like to use this opportunity to explain our proposal in more detail. We think that the character variants that are found in glyph collections such as San-Mitsu-Sho should not be encoded in UCS. The characters variants that we proposed on N4407R occur not only in such glyph collections but also in general documents. In the case of the representation of glyph collections, it is natural to distinguish character variants at the font level through the use of unencoded glyphs. However, in the case of general documents, character variant differences should be treated more carefully. As a result of our survey of general Siddham documents, we concluded that the character variants that we proposed should be distinguished in plain text, meaning that the differences should be preserved at the text encoding level. Of course, encoding such character variants does not prevent them from being handled by the Tenchi-Ji font system.