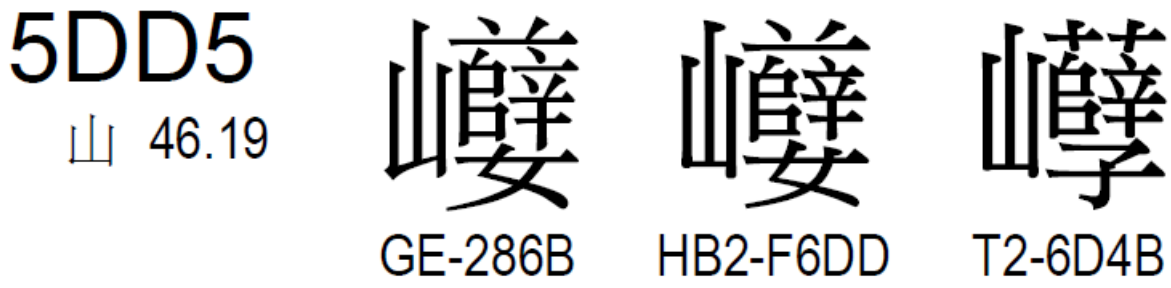
L2/20-020

Subject: Request for reversion of the glyph for T2-6D4B at U+5DD5
Date: 2020-01-09
To: UTC
From: Jaemin Chung
Page: 5

U+5DD5 嶸—a character in the URO (U+4E00–U+9FA5)—seems to be a unification between the two components 女 (G and H) and 子 (T only), which is never acceptable.



I wondered why this kind of huge unification mistake is in the URO, but after checking previous editions of ISO/IEC 10646 and CNS 11643, I now understand what happened.

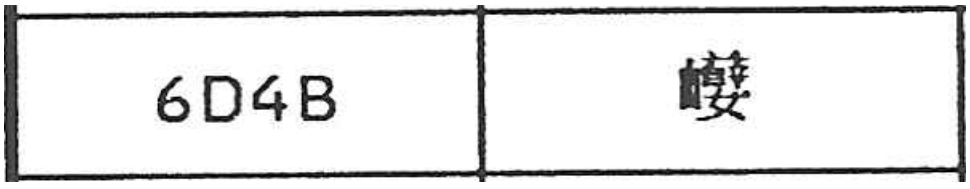
Section 1: What actually happened

Big5 (1984) shows the glyph with the 女 component at 0xF6DD.

Excerpt from Big5

CNS 11643-1986 and CNS 11643-1992 have 嶸 (not 嶂) at 2-6D4B.

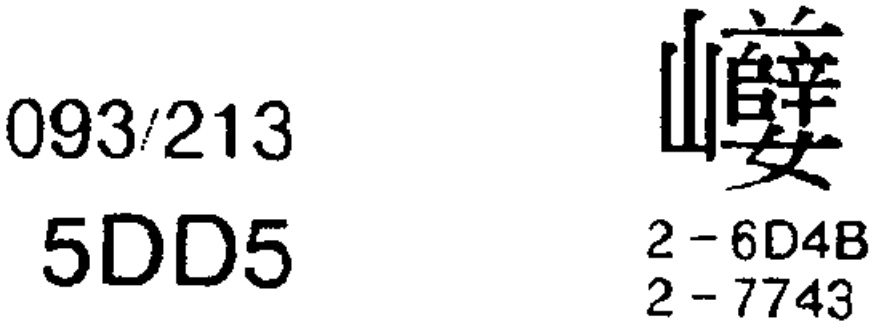
Excerpt from CNS 11643-1986



Excerpt from CNS 11643-1992

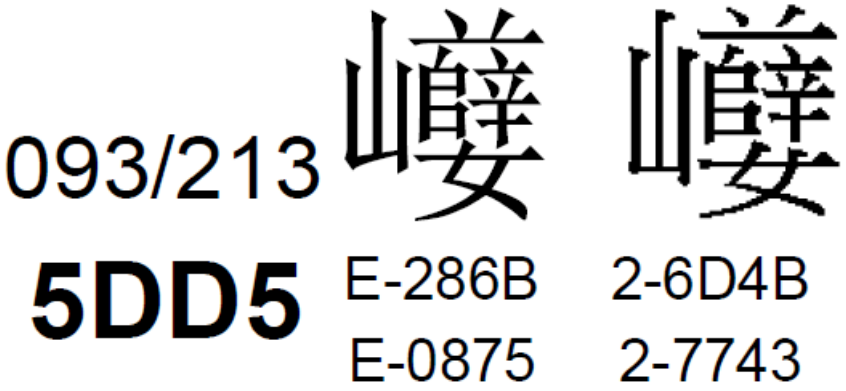
And when the URO was established in the early 1990s, that is what TCA submitted as T2-6D4B.

Excerpt from ISO/IEC 10646-1:1993



In the 2000 edition of ISO/IEC 10646, a G glyph and source reference were added based on the original T glyph.

Excerpt from ISO/IEC 10646-1:2000



But later, TCA silently changed the glyph for T2-6D4B to 嶝 in the 2011 edition (and all the subsequent editions) and causes a problem (and contradicted China who respected the original T glyph).

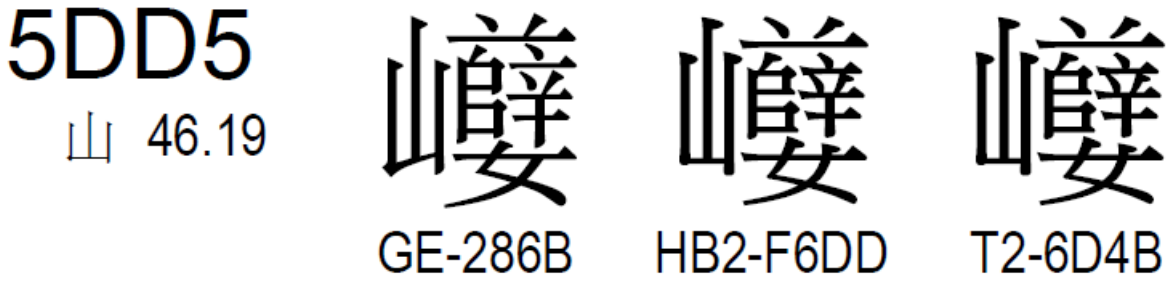
Excerpt from ISO/IEC 10646:2011

Note that the real 巆 is encoded at U+21FD2 巆.

Section 2: Proposed changes

It is very clear that the 巆 glyph cannot stay at U+5DD5 巆, as TCA made a huge non-unifiable change. The T glyph for U+5DD5 巆 must be reverted to 巆.

After this change, the UCS code chart should look something like this:

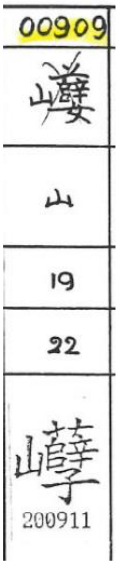


Section 3: Comments, possible issues, and arguments

The following is TCA's response with regard to this issue:
http://appsrv.cse.cuhk.edu.hk/~irg/irg/irg50/IRGN2272R_TCAresponsestoGlyphissue.pdf#page=6

Issue	Code Points and T References	TCA Response
1	5DD5 山 46.19 GE-286B HB2-F6DD T2-6D4B U+5DD5 T2-6D4B	U+5DD5 is in the range of Big5 code, and Microsoft fonts(ex, Microsoft JhengHei, DFKai-SB) are based on this shape □山孌. Although the glyph in Super CJK v14 is □山孌, but *Master Copy of Standard Song Typeface for Chinese Characters* (《國字標準字體宋體母稿》)published by the MOE in 1994 has changed this shape to □山孿. In addition, T2-6D48  is one of the *Table of less frequently used standard Chinese characters* published by the MOE. So, TCA won't change the glyph of T2-6D48, and keep the current code chart.
	21FD2 山 46.20 UCS2003 G4K U+21FD2	

Fig. 6.1 T2-6D4B Glyph on the *Master Copy of Standard Song Typeface for Chinese Characters* (《國字標準字體宋體母稿》),1994.



(Document continued on the next page)

However, this is not a valid reason to have the 嶂 glyph at U+5DD5 巆.

I understand that the Ministry of Education of Taiwan stated in 1994 that 巆 is wrong and 嶂 should be used instead. However, this does NOT mean that 2-6D4B in CNS 11643—which was already submitted to ISO/IEC 10646 before Taiwan MoE's statement—should be altered. Instead, this simply means that a new character needs to be added at a new code position in CNS 11643 (and ISO/IEC 10646).

Making a huge non-unifiable change to a character that is already encoded in ISO/IEC 10646 is not the way to have the needed character. This only causes a problem and pollutes the UCS code chart.

TCA also said that some fonts are already using the 嶂 glyph for U+5DD5 巆. So what? This is because some font developers, for better or for worse, have only the code charts as their sole glyph reference, which unfortunately results in propagating such errors. Since TCA is the culprit, TCA must accept the reversion of the T glyph for U+5DD5 巆 in the UCS code chart.

TCA might complain that their needed character is outside the BMP. If so, what TCA should have done is to propose to encode 嶂 in the BMP before it is encoded in Extension B. Since the real 嶂 is already given a UCS code point, complaining about having a non-BMP character is meaningless.

(End of document)