

Universal Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N2893

Date: 2025-10-15

Source:	China
Author:	Tao Yang (陶扬, CESI) Wang Xieyang (王谢杨, Center for Toponym Research of SISU) Liu Huidan(刘汇丹, Institute of Software, Chinese Academy of Sciences) Su Ruixin(苏瑞欣, Zhonghua Book Company)
Title:	China's advice on the encoding work of CJK Hybrid Characters
Meeting:	IRG #65
Status:	Member's submission
Actions required:	To be considered by IRG
Distribution:	IRG
Medium:	Electronic
Page:	2
Appendix:	None

China holds the view that, in general, it is unsuitable to refer to the following two types of characters as "ideographs": one being hybrid characters composed of Han ideographs together with other scripts or symbols, and the other being characters that consist solely of other scripts or symbols but can be used in running CJK texts.

China insists that the official name for such kind of characters should be "CJK Hybrid Characters" rather than "CJK Hybrid Ideographs". It is recommended to establish a new block named "CJK Hybrid Characters" for encoding the CJK hybrid characters used in CJK texts.

- Given that the font style and application scenarios of Han-Hangul hybrid characters are basically consistent with those of Han ideographs, and there are no obvious non-Han ideograph components in such hybrid characters, we believe that Han-Hangul hybrid characters can continue being encoded in the CJKUI.
- Although Zhuyin Fuhao (Bopomofo), Hiragana, and Katakana originated from Han ideographs, they have developed unique application scenarios and font styles. Moreover, some of the components they contain have never appeared in Han ideographs throughout history, such as ° (U+309C) and 𐄂 (U+31B1). Based on this, we argue that Han-Bopomofo hybrid characters and Han-Kana hybrid characters should not be encoded in CJKUI.

- Hybrid characters composed of Han ideographs with Latin letters, other scripts, or symbols, as well as characters that are made up solely of Kana, Bopomofo, Latin letters, other scripts, or symbols and can be used in running CJK texts, should not be encoded in CJKUI.

China holds the view that, when submitters submit CJK Hybrid Characters, they shall first submit the characters and supporting information to the IRG for discussion. The IRG will determine whether the characters should be encoded and which block they should be encoded into. Subsequently, the submitters shall submit the characters that the IRG has unanimously agreed to include in CJKUI to the next IRG working set. For the characters that the IRG deems inappropriate to be included in CJKUI, the submitters shall submit them to the WG2 for encoding in the "CJK Hybrid Characters" block.

中国认为，总体而言，将由汉字与其他文种、非文字符号构成的混合体字符，或仅由其他文种、非文字符号构成的可混用于中日韩文本的字符称为表意文字是不恰当的。

中国主张，此类字符的正式名称应为“中日韩混合体字符”，而非“中日韩混合体表意文字”。中国建议设立一个新区块，命名为“中日韩混合体字符”，用于编码在中日韩文本中使用的中日韩混合体文字。

- 鉴于汉字-谚文混合体文字的字形风格与应用场景均与汉字基本一致，且该类混合体文字中不存在明显的非汉字部件，因此我们认为可继续在中日韩统一表意文字中编码汉字-谚文混合体文字。

- 注音符、平假名、片假名等虽起源于汉字，但已形成独特的使用场景与字形风格，且其包含的部分部件在历史上从未出现于汉字中（例如“ㄆ”与“𠂔”）。基于此，我们认为汉字-注音符混合体文字、汉字-假名混合体文字不应编入中日韩统一表意文字。

- 由汉字与拉丁字母及其他文字、符号构成的混合体文字，以及仅由假名、注音符、拉丁字母或其他文字、符号构成的可混用于中日韩文本的文字，不应编入中日韩统一表意文字。

中国认为，提交者在提交中日韩混合体字符时，应先将相关字符及辅助信息提交至 IRG 讨论，由 IRG 确定相关字符是否应进行编码及编入哪个区块。后续，提交方需将 IRG 一致同意编入中日韩统一表意文字的字符提交至下一 IRG 工作集；对于 IRG 认为不宜编入中日韩统一表意文字的字符，则提交至 WG2，并将其编码至“中日韩混合体字符”区块。

(EOD)