

Universal Multiple - Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N2731

Date: 2024-10-10

Source:	China
Author:	Li Guoying, Liu Huidan, TAO Yang
Title:	China's Response to IRGN2717
Meeting:	IRG #63
Status:	Member's submission
Actions required:	To be considered by IRG
Distribution:	IRG
Medium:	Electronic
Page:	3
Appendix:	N/A

IRG N2637 and IRGN2717 are discussing the feasibility of encoding script-hybrid ideographs, which involves many issues about IRG's scope of work, CJK's character inclusion category, and the definition of Chinese characters. China believes that these issues should be treated with caution, and puts forward the following opinions on several points that have not been clarified in the two documents.

IRG N2637 和 IRGN2717 这两份文件的讨论内容是汉字-表音部件混合体文字编码的可行性，这涉及到 IRG 的工作范围、CJK 的收录范畴、汉字的定义等诸多问题，中国认为应当审慎对待，对两份文件中没有展开厘清的几个要点，中国提出以下看法。

1 The script-hybrid ideographs should not be encoded as CJKUIs

1 汉字-表音文字混合体文字不应被编入中日韩统一表意文字

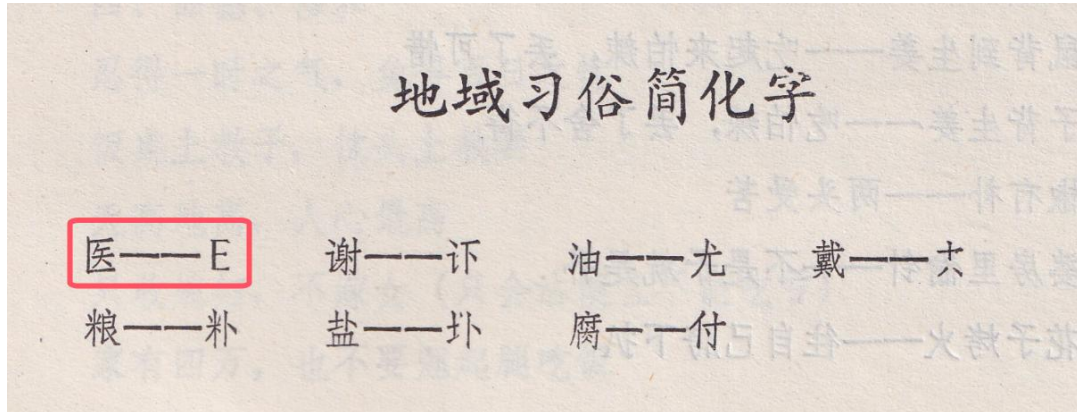
Chinese experts believe that although these script-hybrid characters have phonetic scripts as components, they undoubtedly belong to ideographic scripts. Nevertheless, these characters exceed the general understanding and definition of Chinese characters, and go beyond the scope of technical processing of Chinese characters. If the script-hybrid characters are encoded as CJKUIs, it will bring unnecessary trouble to standard setters and manufacturers.

中国专家认为，这些汉字-表音文字混合体文字虽然毫无疑问属于表意文字，但这类字符超出了对汉字的一般理解和定义，超出了汉字类字符的技术处理范畴。如果将这些文字作为汉字编码，会为标准制定方和厂商带来不必要的麻烦。

Moreover, IRG is an organization that encodes CJKUIs, and its PnP is designed for encoding CJKUIs. The PnP is not suitable for encoding script-hybrid characters such as characters composed by CJKUIs and English letters, CJKUIs and kana, and CJKUIs and hangul. To include such script-hybrid characters in CJKUI, extensive modifications need to be made to IRG PnP, which is very difficult. In addition, encoding such characters together with CJKUIs may confuse the boundaries between letters and Chinese characters, seriously interfering with the existing theories of Chinese

characters. For example, the nature of letters which are used to replace Chinese characters may become controversial if we do so:

而且，IRG 是编码汉字的组织，其 PnP 是为编码汉字设定的，不适应汉字-英文字母、汉字-假名、汉字-谚文等混合体字的编码，将此类混合体字纳入中日韩统一表意文字，需要对 IRG PnP 进行大量修改，相关的工作会很困难。另外，将此类字符与中日韩统一表意文字放在一起编码，可能会混淆字母与汉字的界限，与现有汉字理论产生严重冲突。比如，用于代替汉字的字母的性质可能会产生争议：



2 Agree with encoding script-hybrid ideographs in a new block on Plane 1 in principle

2 原则上同意在第 1 平面设置一个用于编码汉字-表音文字混合体文字的新区块

The UTC brought out in its document that a new block can be set up on plane 1 to encode script-hybrid ideographs. China agrees in principle and recommends that all similar script-hybrid ideographs used in the Cultural Circle of Chinese Characters be encoded in this block. The naming of this block should reflect its relationship with CJKUIs.

统一码委员会在其文件中指出，可在第 1 平面设置单独区域放置混合体字符，中国原则上同意，并建议凡属汉字文化圈的类似方块字符都在该区域处理。该区块的定名应体现其与中日韩统一表意文字的关系。

3 We recommended that script-hybrid ideographs be reviewed by IRG

3 我们建议将 IRG 引入汉字-表音文字混合体文字的审核流程

Some script-hybrid ideographs have very similar shapes to the CJKUIs, which makes them possible to be encoded as CJKUIs. We recommend that IRG decide whether they should be encoded as CJKUIs or be encoded in the special block for script-hybrid ideographs. As IRG is responsible for reviewing whether characters are CJKUIs, reviewing the ideographs falls within its original scope of responsibility.

有一些汉字-表音文字混合体文字的字形与一般的中日韩统一表意文字形体相近，有作为普通汉字编码的可能。我们推荐由 IRG 来决定将这种文字编码为中日韩统一文字或是编入第 1 平面的特别区块。因 IRG 有审查字符是否为汉字的责任，所以 IRG 承担该工作属于其原本的职责范围。

4 Our concerns about encoding script-hybrid ideographs

4 关于编码汉字-表音文字混合体文字的担忧

The Chinese expert group believes that there is a general lack of understanding and study among linguists and coding experts on script-hybrid ideographs, and a consensus has not yet been formed to determine their standardized form, radicals, strokes, variants, and unification rules. So it is too risky to encode them for the time being, and the encoding work should not be carried out too quickly. We suggest that the encoding work should be carried out after there is a proposal that clearly illustrate their history, current use status, and provide methods agreed upon by experts to determine its standardized forms, radicals, strokes, variant forms, and unification rules.

中国专家团认为，文字学和编码专家对汉字-表音文字混合体文字尚普遍缺乏认知和研究，确定其规范字形、部首、笔画、异体字及统合方式的共识尚未形成，现在直接为其编码过于冒险，相关的编码工作不宜过快进行。我们建议相关的编码工作应当在有提案能够清晰论述其历史、应用现状，能够给出专家一致同意的确定其规范字形、部首、笔画、异体字及统合方式等的方法后进行。

(EoD)