

*Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации*

Doc Type: Working Group Document

Title: IVD Registration Group (IRG)

Source: Richard Cook (UC Berkeley), Ken Lunde (Adobe Systems)

Status: Individual Contribution

Action: For consideration by JTC1/SC2/WG2 and UTC

Date: 2011-04-05 (Revised: 2011-04-14)

The IRG (Ideographic Rapporteur Group) and its predecessor the CJK-JRG (CJK Joint Research Group) worked within WG2 to resolve issues relating to unification of the various Han (CJK) national encoding standards. Those original (pre-unification) standards are now obsolete (legacy encodings), but their successors continue to drive CJK encoding processes. CJK characters, code charts and property tables constitute by far the largest UCS components, and yet are best viewed as only a formulation of the complex encoding problems that continue to this day. Rules of thumb describing Han unification are outlined in ISO/IEC 10646 *Annex S* (informative). Comprehensive revision of that document was recently attempted, in recognition of the need for a working encoding model. This work succeeded only in clarifying some of the issues, but did not result in a substantial revision.

It now seems clear that the vast majority of abstract Han characters have already been encoded, in some cases multiple times. Most forms now under consideration for future encoding are identifiable as useless variants of encoded abstract characters. The encoding of such forms serves no practical purpose, degrades the international (and national) standards, imposes high costs on implementers, and confuses end-users.

For these reasons, it is recommended that further revision of *Annex S* not be attempted. Rather, *Annex S* should be frozen in its current state, and framed as a largely obsolete record of the original unification processes. IRG should be re-purposed, to serve within WG2 as the “IVD Registration Group”, to focus on processing IVD (Ideographic Variation Database) registration requests (UAX #37). IRG would contribute to the development of UAX #37, to codify guidelines for distinguishing CJK variants. IRG might rarely find that a form submitted for IVD registration seems more like an abstract character (and less like a unifiable variant), in which case the group could assist the registrant in the preparation of a proposal to encode a new abstract character.