

Universal Multiple-Octet Coded Character Set  
International Organization for Standardization

**Doc Type:** Working Group Document

**Title:** Proposal to establish a CJK Unified Ideographs “Urgently Needed Characters” process

**Authors:** Dr. Ken Lunde (小林劍) & John Jenkins (井作恆)

**Source:** The Unicode Consortium & US National Body

**Status:** National Body Contribution

**Action:** For consideration by JTC1/SC2/WG2

**Date:** 2012-02-10 (originally submitted to the UTC on 2011-11-17, and revised on 2012-02-07)

## Background

The process of standardizing repertoires of CJK Unified Ideographs is long and cumbersome, and is almost always measured in years. This is primarily because the typical CJK Unified Ideograph repertoire includes thousands or tens of thousands of characters, and thus requires several rounds of review and discussion before it can be standardized. Extension E, for example, began as Extension D, whose national body submissions were accepted in early 2007, and incorporated characters that were deferred from Extension C. Extension E is currently at the final stages of standardization.

To address this particular process shortcoming, the IRG established a one-time UNC (*Urgently Needed Characters*) repertoire as one of the IRG 29 Resolutions (see IRG N1377, specifically Resolution IRG M29.5), which eventually became CJK Unified Ideographs Extension D, with 222 CJK Unified Ideographs (U+2B740 through U+2B81D), and which was subsequently included in ISO/IEC 10646:2003 Amendment 8, and in Unicode Version 6.0.

Without a formalized UNC-like process in place, which would serve as a parallel pipeline for smaller repertoires of urgently-needed CJK Unified Ideographs, it is extraordinarily difficult for national bodies to standardize smaller sets of urgently-needed CJK Unified Ideographs in a timely manner.

## Proposal

We hereby propose that a formalized CJK Unified Ideographs “Urgently Needed Characters” process be established. In order for such a process to function effectively, the standardization timeline must be *short*, and the repertoires must be *small*.

To address the timeline issue, in that it must be short, we recommend that each repertoire be standardized on a regular basis, such as every one or two years, and subsequently included in an amendment or new edition of the standard.

To address the size issue of the repertoire, in that it must be small, we recommend that each national body be limited to 25 character submissions per repertoire, in terms of what is originally submitted, not what actually becomes standardized. In other words, national bodies should not be allowed to submit additional characters to replace ones that are removed for reasons such as lack of sufficient evidence or unification. Any exceptions to this 25-character limit must be approved by the IRG, and the total number of submissions per repertoire for all national bodies combined should not exceed 200. With a “25 characters per national body per repertoire” submission limit in place, the effort needed to bring each repertoire to a state in which it can be standardized becomes much more manageable than for the larger repertoires.

We request that WG2 direct the IRG to formalize an “Urgently Needed Characters” process—as a parallel pipeline to that intended for standard CJK Unified Ideograph submissions—per the above recommendations, and that the IRG be further tasked with documenting the detailed procedures in their Principles and Procedures document.