

Title: Unicode Liaison Report to WG2

Date: 2014-2-26

Source: Unicode Consortium

Status: Liaison contribution

Action: For review by WG2 experts

Distribution: WG2

The Unicode Consortium is pleased to report on-going progress in development of the Universal Character Set resulting from collaboration with SC2, as well as progress on the Unicode Standard and related standards and technologies.

Preparation of Unicode 7.0 for publication

Work has been in progress on version 7.0 of the Unicode Standard, with a target date for publication in July 2014.

The main focus for version 7.0 is the addition of new characters and catching up with progress in that regard in ISO/IEC 10646. It will synchronize with ISO/IEC 10646:2012 plus amendments 1 and 2. Additionally, it will fast-track inclusion of U+20BD RUBLE SIGN.

Preliminary content for Unicode 7.0 has been made publicly available for review and feedback. For additional information, see <http://www.unicode.org/review/pri271/>.

Unicode Technical Reports normatively referenced in ISO/IEC 10646

It is understood that ISO/IEC 10646 makes normative reference to these three specifications maintained by the Unicode Consortium:

- UAX #9 Unicode Bidirectional Algorithm
- UAX #15 Unicode Normalization Forms
- UTS #37 Ideographic Variation Database

Current versions for each of these are as follows:

- UAX #9, UAX #15: versions published as part of Unicode 6.3, September 2013. The 6.3 version of UAX #9 included substantial changes; see the previous liaison report ([N4447](#)) for details.
- UTS #37: version 3.7, published November 7, 2011.

Updates to UAX #9 and UAX #15 are being prepared for Unicode 7.0; however, there are no substantive changes to report. (For this reason, drafts are not being submitted to the WG2 document register for review.)

Software vendors that have begun to implement the Unicode 6.3 version of the Unicode Bidirectional Algorithm have identified some details in the revised specification that require revision. These will be

prepared for a future version of the Unicode Standard to appear in 2015. WG2 members can expect a draft update to UAX #9 to be submitted for WG2 review in Q1 of 2015.

Publication Schedule for the Unicode Standard

In the past, the Unicode Standard has not been published on any fixed schedule—in the past, releases have ranged from being 3 months apart to almost a year and a half apart. The Unicode Consortium is now involved in publishing several products that get updated on a somewhat frequent basis, and in order to manage workload is moving to a regular publication schedule for the two largest product, CLDR and The Unicode Standard.

As a result, the publication for The Unicode Standard going forward will be to publish a new version every summer. Unicode 7.0 will be released in July 2014; the next version after that, in summer 2015; and so on.

This publication schedule will have certain implications for synchronization with ISO/IEC 10646. The publication process for Unicode will require that the character repertoire be finalized by late January for a version to be released the following summer. In terms of process for 10646, characters can be included in a given version of Unicode if, by late January of that year, the characters are in a published edition or amendment of 10646 or are at least ready for approval stage (i.e., a DAM or DIS ballot has been completed and ballot comments have been resolved).

For example, if WG2 62 determines that a repertoire for Amendment 1 is ready for a DAM ballot, and that DAM ballot has passed after ballot comments are resolved in September at WG 63 in Colombo, then it will be possible to include that repertoire in Unicode's 2015 release. But, if work is started at WG2 62 on Amendment 2 with a DAM prepared after Colombo, that repertoire could not be considered for Unicode's 2015 release since it would not be stabilized by January 2015.

WG2 may want to take Unicode's schedule into consideration when planning timelines for new editions and amendments. In particular, it will aide in synchronization if ballot resolution meetings for enquiry drafts (DIS, DAM) can be timed for Q4 of any given calendar year, or mid-January at the latest.

Accessing Draft Documents

JTC1/SC2 was recently required to migrate the SC2 document register into the ISOTC online system. This has meant that access control to SC2 documents is now managed by ISO/IEC administration, whereas it used to be managed by the SC2 secretariat. Drafts of new editions or amendments to ISO/IEC 10646, as a result, are no longer accessible for WG2 members. Some members may have access to these documents via their national body, though this is inconsistent:

- WG2 experts that have a relationship with some NB may have different experience in how reliably their NB relays documents. In some cases, documents might not be relayed or might not be relayed in a timely manner; in some cases, mix-ups in documents have been known to occur.
- Some WG2 experts (including some contributing editors) might not have a direct relationship with any NB.

Most recently, the Unicode Consortium has not been able to obtain the ballot documents for the DIS of the 4th edition of 10646. Other WG2 members may also have faced this. This has created an obstacle to

successful participation in WG2: it is difficult for a given expert to discuss resolution of comments on the DIS when it they have no way to look at the content of the DIS.

This may be seen as a problem for members to take up with their NB, or as a problem for SC2. In as much as it has impacted the work WG2, though, it is also a WG2 problem.

For this reason, we suggest that WG2 take a resolution to communicate a concern up to SC2. The following is draft text for a resolution that might be considered:

Whereas access to enquiry drafts are now controlled via the ISO/IEC “OpenText” content server, with many WG2 members not having sufficient rights to access these drafts in the JTC1/SC2 N-documents list on that system; and whereas this results in WG2 members not having access to documents that need to be discussed in ballot resolution meeting,

WG2 requests that SC2 and its secretariat can explore options to enable WG2 members to have access to enquiry draft documents for review and to facilitate consideration of these drafts at WG2 meetings.

Common Locale Data Repository (CLDR)

Unicode CLDR, Version 24, was released on September 18, 2013. That version contains data for 238 languages and 259 territories—740 locales in all. The next version is being prepared for release in March, 2014.

The Unicode Consortium feels confident that National Bodies and experts represented in WG2 will find the CLDR offers useful benefits in enabling support in software products for languages and cultures from across the world. As always, experts in WG2 are invited to participate in the on-going development of CLDR. Current information on CLDR can be found on the Unicode Web site at <http://cldr.unicode.org/>.