



Date: 2007-09-20

**ISO/IEC JTC 1/SC 2/WG 2**  
**Universal Multiple-Octet Coded Character Set (UCS) - ISO/IEC 10646**  
 Secretariat: ANSI

<b>Doc Type</b>	Liaison Report
<b>Title:</b>	Unicode Liaison Report – meeting 51
<b>Source:</b>	Peter Constable—Unicode Liaison
<b>Action:</b>	None
<b>Distribution:</b>	SC2/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.

#### **Unicode Technical Committee Review of WG2 #50 (Frankfurt)**

Various actions were taken by WG2 in Frankfurt in relation to FDAM 3, FPDAM 4 and PDAM 5 that impacted prior decisions made by UTC in relation future versions of the Unicode Standard. WG2 also took actions adding new characters to one of these amendments which had not previously approved by UTC for addition to Unicode. UTC reviewed the WG2 actions and, except as noted below, took actions to implement corresponding changes in Unicode so as to maintain synchronization.

WG2 approved the addition of Avestan script to PDAM5 at the Frankfurt meeting. UTC had previously approved the encoding in Unicode of Avestan script as described in WG2/N3197 with the exception of one character: 10B38 AVESTAN SEPARATION POINT. In review of WG2 #50, UTC considered this character further. UTC experts were concerned with the proliferation of script-specific word-separator-dot characters that differ from script to script only in minor differences in appearance. After consideration of issues, UTC concluded that it would be better to encode one generic (non-script-specific) word-separator dot character that could be used for multiple scripts with the assumption that the range of acceptable glyph representations would allow for the differences in appearance in different scripts. Accordingly, UTC recommends to WG2 that 10B38 AVESTAN SEPARATION POINT be withdrawn from Amendment 5 and that a generic character WORD SEPARATOR MIDDLE DOT be encoded in its place.

In all other respects, UTC endorses the actions taken by WG2 #50 in relation to FDAM 3, FPDAM 4 and PDAM 5.

### **New Edition of ISO/IEC 10646 (WG2/N3274)**

UTC was given a presentation by Michel Suignard on work to prepare a new edition of ISO/IEC 10646, and is pleased to see work that is being done by WG2 to bring ISO/IEC 10646 and the Unicode Standard into even closer alignment. UTC believes that the terminology changes under consideration will be beneficial for implementations of ISO/IEC 10646 and improve the overall usefulness of that standard.

### **Unicode 5.1**

UTC anticipates the publication of version 5.1 of the Unicode Standard in 2008, to be synchronized with ISO/IEC 10646:2003 plus Amendment 3 and Amendment 4. UTC will review the actions of WG2 #51 (Hangzhou) in relation to FDAM 4 at their next meeting and determine what actions are needed to ensure synchronization of Unicode 5.1 with Amendment 3 and FDAM 4.

### **Common Locale Data Repository (CLDR) Version 1.5**

The Unicode Consortium released version 1.5 of CLDR in July 2007. Version 1.5 contains 42% more locale data than the previous release, with over 27,000 new or modified data items entered by over 160 different contributors. Major contributors include several software vendors and also official representatives from a number of countries. This version contains data for 394 locales, covering 135 languages and 149 territories.

The Unicode Consortium feels confident that National Bodies and experts represented in WG2 will find the new version of 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://unicode.org/cldr/>.