

Subject: UTC Item: Additional proposed Improvement for Section 4.6

L2/10-276

From: Asmus Freytag <asmusf@ix.netcom.com>

Date: Fri, 30 Jul 2010 15:49:09 -0700

To: Mistress of the Docs <l2doc@unicode.org>, Lisa Moore <lisam@us.ibm.com>

The following are editorial improvements suggested for Section 4.6, going beyond what is currently covered in the proposals in document L2/10-262

Problem:

Description of much of the numeric use of characters is scattered across the standard.

Solution:

Create one place where links to existing descriptions are collected.

I suggest that a rather complete summary of all exceptional cases of numerical character use be maintained in a central place (section 4.6 comes to mind). This would provide implementers with the information they need to track down unusual behavior.

Currently, the information is not easily locatable, other than scanning all chapters or searching for keywords (which may not be selective enough or not applied with enough consistency). By its nature, the information is useful to implementers of cross-script (or script neutral) implementations.

A table of links (references) is sufficient for this purpose - there's no need to duplicate any description or add a lot of explanatory text.

The table could be broken down into two main sections.

The first section would link to descriptions of unusual behavior for characters used as decimal-radix digits.

Examples of this are:

- a) Arabic using two complete series of digits
- b) New Thai Lue using an extra digit 1
- c) Han digits being scattered and used in two different types of numeric expressions
- d) ASCII digits being used for some scripts as preferred decimal-radix digits, because their native number system is not, or not exclusively decimal-radix (Han, Ethiopic..)
- e) Compatibility characters based on decimal digits

A separate section of the table should pull together links to all the descriptions of non-decimal radix number systems

that are discussed in the Standard, such as

- a) roman numerals
- b) Greek characters used as numerals
- e) Ethiopic numerals
- c)
