

ISO/IEC JTC 1/SC 2 N 3804

DATE: 2005-08-03

REPLACES: SC 2N3799

ISO/IEC JTC 1/SC 2
Coded Character Sets
Secretariat: **Japan (JISC)**

L2/05-221

DOC. TYPE	Summary of Voting/Table of Replies						
TITLE	Revised Summary of Voting on SC 2 N 3792: ISO/IEC 14651/PDAM 3, Information technology -- International string ordering and comparison -- Method for comparing character strings and description of the common template tailorable ordering -- AMENDMENT 3						
SOURCE	SC 2 Secretariat						
PROJECT	JTC 1.02.14651.00.03						
STATUS	This document is forwarded to Editing Group on 14651 for resolution of comments. The Editing group is requested to prepare a disposition of comments report, revised text and a recommendation for further processing.						
ACTION ID	FYI						
DUE DATE							
DISTRIBUTION	P, O and L Members of ISO/IEC JTC 1/SC 2 ; ISO/IEC JTC 1 Secretariat; ISO/IEC ITTF						
ACCESS LEVEL	Def						
ISSUE NO.	220						
FILE	<table border="1"><tr><td>NAME</td><td>02n3804.pdf</td></tr><tr><td>SIZE (KB)</td><td></td></tr><tr><td>PAGES</td><td>6</td></tr></table>	NAME	02n3804.pdf	SIZE (KB)		PAGES	6
NAME	02n3804.pdf						
SIZE (KB)							
PAGES	6						

Secretariat ISO/IEC JTC 1/SC 2 - IPSJ/ITSCJ *(Information Processing Society of Japan/Information Technology Standards Commission of Japan) Room 308-3, Kikai-Shinko-Kaikan Bldg., 3-5-8, Shiba-Koen, Minato-ku, Tokyo 105-0011 Japan *Standard Organization Accredited by JISC
Telephone: +81-3-3431-2808; Facsimile: +81-3-3431-6493; E-mail: kimura@itscj.ipsj.or.jp

**Summary of Voting on
ISO/IEC JTC 1/SC 2 N 3792 :
ISO/IEC 14651/PDAM. 3, Information technology -- International string ordering and comparison -- Method
for comparing character strings and description of the common template tailorable ordering -- AMENDMENT
3**

Q1 : PDAM

	Q1			Not yet voted	Comments Reasons
	Approve	Disapprove	Abstention		
P-Member					
Armenia				X	
Belgium				X	
Canada	X				
China	X				
Denmark			X		
Egypt				X	
Finland	X				
France			X		
Germany	X				
Greece				X	
Iceland				X	
India				X	
Iran, Islamic Republic of	X				
Ireland		X			See Attachment
Italy	X				
Japan			X		
Korea, Democratic People's Republic				X	
Korea, Republic of	X				
Mongolia				X	
Morocco		X			See Attachment
Norway	X				
Poland				X	
Romania				X	
Russian Federation				X	
Serbia and Montenegro				X	
Singapore	X				
Slovakia				X	
Sweden	X				
Thailand		X			
Tunisia				X	
USA	X				See Attachment
Total (31)	11	3	3	14	

*: Approve with comments

#: Acceptance of the reasons and appropriate changes in the text will change the vote to approval.

Ireland disapproves the draft with the technical comment given below. Acceptance of the comment and appropriate changes to the text will change our vote to approval.

T1. We believe that the tailorings of Thai and Lao logical reordering are burdensome to the end user, since everyone who wants these scripts to order will want correct ordering anyway. We believe that the Thai and Lao logical reordering should be built into the table instead of being a tailoring to facilitate that.

Moroccan vote to ISO/IEC JTC 1/SC 2 N 3792:

Morocco disapproves the PDAM 3 for ISO/IEC 14651:2001 for the comment given below. Acceptance of this comment will change our vote to approval.

The template table contains a mistake in the following line:

<U2D6F> <S2D61>;<BASE>;<MNN>;<U2D6F> % LETTRE
MODIFICATIVE TIFINAGHE DE LABIALISATION (tamart)

We request to correct it as follows:

Reorder-after S2D65

**<S2D7F> % LETTRE MODIFICATIVE TIFINAGHE DE
LABIALISATION (tamart)**

**<U2D6F> <S2D7F>;<BASE>;<MNN>;<U2D6F> % LETTRE
MODIFICATIVE TIFINAGHE DE LABIALISATION (tamart)**

The US National Body votes to approve with comments PDAM ballot for ISO/IEC 14651/PDAM3, Information technology -- International string ordering and comparison -- Method for comparing character strings and description of the common template tailorable ordering -- AMENDMENT 3, SC2 N3792.

Technical

1. CTT

The U.S. suggests that the CTT used in the balloted amendment (based on the file ctt14651-4.1.0d6.txt dated April 7, 2005 submitted to the editor) be updated to a slightly later revision (ctt14651-4.1.0d11.txt dated May 2, 2005) provided to the editor at the following URL:

<http://www.unicode.org/~whistler/ctt14651-4.1.0d11.txt>

The rationale for this change is twofold:

First, the d11 delta of the CTT table is in exact synchronization with the recently published Unicode 4.1.0 version of Unicode Collation Algorithm, which will likely be the de facto implementation of the ISO 14651 CTT (and its tailorings) on many important platforms. Having the 14651 Amendment and the Unicode Collation Algorithm in close synchronization is important for ensuring interoperability of implementations of collation based on the 14651 standard.

Second, the main change between the d6 delta of the CTT table and the d11 delta is the introduction of predefined weightings for Thai and Lao consonant/vowel combinations that involve logical reordering. It was intended from early on that these changes to the CTT table be included in the balloted table, but they were omitted because of technical difficulties in production of the table under the tight time constraints to produce the ballot. The predefined weightings for Thai and Lao consonant/vowel combinations result in a simpler-to-implement table to produce correct Thai or Lao string ordering, with no requirement of complex tailoring or prehandling of data to produce correct results. This change to the table was requested some time ago by Sweden (among others), to simplify the task of producing correct string ordering for Thai. The U.S. NB concurs with that request, and feels that the introduction of the combination weightings as specified in ctt14651-4.1.0d11.txt will result in better handling of Thai or Lao data in most implementations.

Furthermore, this change to the CTT is in essential agreement with the rationale for adding section B.5 in Annex B -- the proposed text for a tailoring for Khmer, Thai, and Lao, as spelled out in the Amendment text. However, in our opinion, the Thai and Lao portions of that addition are now better handled by the CTT itself, instead of in the informative suggested tailoring in section B.5.

2. References

Because of the importance of implementation synchronization, the U.S. requests that full, formal references to both Version 4.1 of the Unicode Standard and Version 4.1

of the Unicode Collation Algorithm (UTS #10) be added to the references for the Amendment.

3. Annex B, Section B.5 Tailoring Addition

Although Annex B.5 is informative, the U.S. feels that it is important that the suggested tailoring there (as shown in the ballot text) be correctly articulated with the CTT itself. Because we are suggesting that the Thai and Lao reordering dealt with in the suggested text for Example 5 in section B.5 be directly addressed in the CTT (in delta 11: ctt14651-4.1.0d11.txt), the relevant sections in Example 5 dealing with Thai and Lao should then be removed, as they would not be required for handling of Thai and Lao data.

In particular, all of the Thai and Lao lines (both collating-element definitions and actual weighting lines) should be removed from Example 5. Also, in the introductory text paragraph and title references to "Khmer, Thai, and Lao" should be emended to simply "Khmer", and the text beginning "Secondly, the 'swapping' for Thai and Lao..." to the end of the paragraph should be stricken.

4. Annex C, Section C.2.2 Vowel/consonant rearrangement

The U.S. suggests editorial revision of the initial paragraph to accord with the above changes requested for the CTT and Example 5.

Revise the first sentence to read:

"Regarding the handling of pre-vowels, either a collation preparation or collating-element grouping is required."

Revise the fourth sentence to read:

"Another way to manage this is by means of collating-element formation -- the approach taken both by the CTT of this standard and by the collation weighting table of the Unicode Collation Algorithm (UTS #10)."

Editorial

The last sentence in the first paragraph of C.2.2 has a typo:

"should be handles just like" --> "should be handled just like"