



ISO/IEC JTC 1/SC 2/WG 2 **N3103**  
DATE: 2006-08-25

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

**DOC TYPE:** Meeting Minutes

**TITLE:** Unconfirmed minutes of WG 2 meeting 48  
Microsoft Campus, Mountain View, CA, USA; 2006-04-24/27

**SOURCE:** V.S. Umamaheswaran, Recording Secretary, and Mike Ksar, Convener

**PROJECT:** JTC 1.02.18 – ISO/IEC 10646

**STATUS:** SC 2/WG 2 participants are requested to review the attached unconfirmed minutes, act on appropriate noted action items, and to send any comments or corrections to the convener as soon as possible but no later than 2006-09-15.

**ACTION ID:** ACT

**DUE DATE:** 2006-09-15

**DISTRIBUTION:** SC 2/WG 2 members and Liaison organizations

**MEDIUM:** Acrobat PDF file

**NO. OF PAGES:** 51 (including cover sheet)

Mike Ksar  
Convener – ISO/IEC/JTC 1/SC 2/WG 2  
Microsoft Corporation  
One Microsoft Way  
Bldg 24/2217  
Redmond, WA 98052-6399

Phone: +1 425 707-6973  
Fax: +1 425 936-7329  
email: [mikeksar@microsoft.com](mailto:mikeksar@microsoft.com)  
or [mikeksar@10646.com](mailto:mikeksar@10646.com)

**ISO**  
**International Organization for Standardization**  
**Organisation Internationale de Normalisation**

ISO/IEC JTC 1/SC 2/WG 2  
Universal Multiple-Octet Coded Character Set (UCS)

ISO/IEC JTC 1/SC 2/WG 2 **N3103**

Date: 2006-08-25

<b>Title:</b>	Unconfirmed minutes of WG 2 meeting 48 Microsoft Campus, Mountain View, CA, USA; 2006-04-24/27
<b>Source:</b>	V.S. Umamaheswaran ( <a href="mailto:umavs@ca.ibm.com">umavs@ca.ibm.com</a> ), Recording Secretary Mike Ksar ( <a href="mailto:mikeksar@microsoft.com">mikeksar@microsoft.com</a> ), Convener
<b>Action:</b>	WG 2 members and Liaison organizations
<b>Distribution:</b>	ISO/IEC JTC 1/SC 2/WG 2 members and liaison organizations

## 1 Opening and roll call

Input document:

[3047](#) 2nd Call - Meeting 48; Ksar; 2006-03-09

Mr. Mike Ksar convened the meeting at 10.07h. The Unicode Consortium is our host. Ms. Magda Danish will give us some information about the logistics.

Ms. Magda Danish: I welcome you all to this meeting hosted by the Unicode Consortium. If you need any assistance I will be able to assist you. She went on to explain some of the logistics about the meeting.

Mr. Mike Ksar: We are in building 1 for all this week. Today we are in the Saturn conference room. For the next four days we will be in the Executive conference room. Ms. Magda Danish has distributed a hardcopy and CD containing all the documents relevant for this meeting, one for each national body and hardcopies for the editors and the secretary. Delegates can copy these into their local machines. I received a note from Mr. Chen Zhuang that the Chinese delegates will not be able to attend the meeting. Dr. Lu Qin should be here and be able to present the Chinese contributions.

I have also information from the SC 2 chair that SC 2 meeting may wind up by Thursday evening -- probably have to stay till 7pm on Thursday. Thursday 1 to 3 pm will be the ISO/IEC 14651 editing group meeting.

The agenda is in document [N3055](#).

### 1.1 Roll Call

Input document:

[3001](#) Updated Experts List; Mike Ksar; 2006-02-27

The following **25** delegates representing **6** national bodies, **2** liaison organizations, SC 2 secretariat and **1** guest organization were present at different times during the meeting.

Name	Representing	Affiliation
Alain LaBonté	Canada Editor 14651	Ministère des services gouvernementaux du Québec
V. S. (Uma) Umamaheswaran	Canada; Recording Secretary	IBM Canada Ltd.
Erkki I. Kolehmainen	Finland	Research Institute for the Languages of Finland
U Thein OO	Guest (Myanmar)	Myanmar Computer Federation
Zaw HTUT	Guest (Myanmar)	Myanmar Computer Federation
Michael Everson	Ireland; Contributing Editor	Evertype
Qin LU	IRG Rapporteur	Hong Kong Polytechnic University
Kazuhito OHMAKI	Japan	National Institute of Advanced Industrial Science and Technology
Yasuhiro ANAN	Japan	Microsoft Japan
KIM Hanbitnari	Korea (Republic of)	Korean Language Society
KIM Kyongsok	Korea (Republic of)	Pusan National University
Dae Hyuk AHN	Korea (Republic of)	Microsoft Korea

Name	Representing	Affiliation
Elzbieta Broma-Wrzesień	Poland	Telekomunikacja Polska, Poland
Tatsuo KOBAYASHI	SC 2 Chair, Japan	Justsystem Corporation
Toshiko KIMURA	SC 2 Secretariat	IPSI/ITSCJ
TSENG Shih-Shyeng	TCA	Academia Sinica
WEI Lin-Mei	TCA	Chinese Foundation Of Digitization Technology
Ienup Sung	USA	Sun Microsystems
Deborah Anderson	USA	University Of California At Berkeley
Lisa Moore	USA	IBM Corporation
Eric Muller	USA	Adobe, Inc.
Asmus Freytag	Liaison - The Unicode Consortium; Contributing Editor	Unicode Inc.
Ken Whistler	USA; Contributing Editor	Sybase Inc.
Mike Ksar	USA; Convener	Microsoft Corporation
Michel Suignard	USA; Project Editor	Microsoft Corporation

A printout of document [N3001](#) containing the WG 2 experts list was circulated for delegates to mark any corrections and indicate their attendance at this meeting. Delegates were also requested to give their business cards to the secretary to avoid inaccuracies in the attendance list and in the experts list. The drafting committee consisted of the recording secretary, the convener, the project editor and contributing editors.

## 2 Approval of the agenda

Input document:  
[3055](#) Preliminary Agenda - Meeting 48; Ksar; 2006-04-16

Mr. Mike Ksar went through the agenda in document [N3055](#).

Mr. Erkki Kolehmainen: Please add 'vocabulary group' under JTC 1 matters. Documents [N3048](#) and [N3090](#) on agenda item 8.20 on Lithuanian are on two separate topics. Will the WG 2 report to SC 2 be discussed - FYI only.

Mr. Yasuhiro Anan: Proposal for collection from Japan in document [N3091](#) is to be added under item 9.4. The information for next meeting of WG 2 in Japan is in document [N3093](#). There is an updated document N2985 on Mahjong under item 8.30.

Mr. Michael Everson: Updated documents are available for 8.8 Rejang, 8.11 Mayanist and 8.29 Vai . Currently listed documents will be replaced with updated versions.

Mr. Mike Ksar: There will be an ad hoc on Myanmar. I have requested Mr. Erkki Kohlemainen to be the chair. Dr. Asmus Freytag, Mr. Michael Everson, Dr. Ken Whistler, Ms. Debbie Anderson, Mr. U Thein Oo, Mr. Zaw Htut and Mr. Yasuhiro Anan identified themselves as participants in the ad hoc. Mr. Peter Constable will not attend can be called if needed. The ad hoc can meet during lunch and breaks -- we can extend the lunch till 2 pm if possible. The ad hoc can also meet outside the meeting location. We may be able to find another room after 5 pm -- Mr. Yasuhiro Anan from Microsoft can be the escorting person.

The agenda is organized the way it is -- gives a chance to the editor to address the Amendment 3 specifics separately. All items under item 8 are for discussion and possible additions to either Amendment 3 or a future Amendment.

The agenda was approved with the above updates. Additional changes made during the progress of the meeting are included in the appropriate sections in this document. Some of the agenda items have been reorganized or renumbered in these minutes.

The following table of contents reflect where the items are discussed and recorded.

Item Number	Title	Page
1	Opening and roll call	1
1.1	Roll Call	1
2	Approval of the agenda	2
3	Approval of minutes of meetings 46 and 47	4
4	Review action items	4
4.1	Outstanding action items from meeting 43, 2003-12-09/12, Tokyo, Japan	5
4.2	Outstanding action items from meeting 46, 2005-01-24/28, Xiamen, China	5
4.3	New action items from meeting 47, 2005-09-12/15, Sophia Antipolis, France	5
5	JTC 1 and ITTF matters	12
5.1	ITU SG17 Requirements for future editions	12
5.2	Vocabulary work group	13
5.3	Other JTC 1 documents	13
6	SC 2 matters	13
6.1	SC 2 Program of Work	13
6.2	Ballot results - Amendment 3	13
7	WG 2 matters / housekeeping	14
7.1	Updates to Roadmap per N2526 - Action item 43-9	14
7.2	Restrict ballot comments to content under ballot - reminder	14
7.3	Missing resolution from M47 on Latin phonetic & orthographic characters	14
7.4	Update to Principles & Procedures	15
7.5	WG 2 report to SC 2 plenary	15
8	New and updated contributions – not part of Amendment 3	15
8.1	Tibetan additions	15
8.1.1	Four Tibetan characters for Balti	15
8.1.2	One Tibetan astrological character	15
8.1.3	Three archaic Tibetan characters	16
8.2	Kaithi script	16
8.3	Malayalam additions	16
8.4	Lycian and Lydian scripts	17
8.5	Carian script	17
8.6	Gurmukhi additions	17
8.6.1	Gurmukhi tone mark Udaat	17
8.6.2	Gurmukhi sign Yakash	18
8.7	Sundanese script	18
8.8	Rejang script	18
8.9	Kayah Li script	19
8.10	Medievalist characters	19
8.11	Mayanist Latin letters	20
8.12	Ordbok characters	21
8.13	Addition of two Mathematical brackets	22
8.14	Arabic Math characters	22
8.15	Mongolian – Manchu Ali Gali letter	23
8.16	Myanmar additions	24
8.16.1	Additional 7 Myanmar characters disunified	24
8.16.2	Myanmar additions for Mon and S'gaw Karen languages	25
8.17	Saurashtra script	25
8.17.1	Correction of a Saurashtra character name in PDAM3	25
8.17.2	Implementation of the Saurashtra script in N2969	26
8.18	Lithuanian	26
8.18.1	Named sequences and collection for Lithuanian	26
8.18.2	Two combining characters for Lithuanian phonetics	28
8.19	CJK Strokes additions	28
8.20	Phaistos Disc script	29
8.21	Three Uralic phonetic additions	30
8.22	Support for OI Chiki script	30
8.23	Karen, Shan and Kayah characters (N3080)	31
8.24	Vai additions	31
8.25	Mahjong symbols	31
8.26	Cyrillic combining marks	32
8.27	Proposed additions to Amendment 3	32
8.27.1	Proposals from US national body	32
8.27.1.1	Latin Characters for Egyptian transliteration	32
8.27.1.2	Lao character name annotations	33
8.27.2	Proposals from UK national body	33
8.27.3	Proposals from Irish national body	33
9	New contributions/feedback resulting from action items	34
9.1	Clarification - Hangul Syllables	34
9.2	Improving formal definition for control characters	35
9.3	Ideographic Variation Database	36
9.4	Collection identifiers for Japanese subsets	37
10	Amendment 3	38
10.1	Proposed disposition of comments on Amendment 3	38

Item Number	Title	Page
10.1.1	Ireland: Disapproval	38
10.1.2	Japan, Yes with comments	39
10.1.3	United Kingdom: Yes with comments:	39
10.1.4	USA: Yes with comments	39
10.1.5	Final disposition	39
10.2	Progression of Amendment 3	39
11	Publication issues	40
12	IRG status and reports	40
13	Defect reports	41
14	Liaison reports	41
14.1	The Unicode Consortium	41
14.2	IETF	42
14.3	JTC 1/SC 22	42
14.4	W3C	42
14.5	UC Berkeley - request for liaison	42
15	Other business	42
15.1	Web Site	42
15.2	Future meetings	43
15.3	Closing	43
15.3.1	Approval of resolutions of meeting 48	43
15.4	Adjournment	44
16	Action Items	44
16.1	Outstanding action items from meeting 47, 2005-09-12/15, Sophia Antipolis, France	44
16.2	New action items from meeting 48, 2006-04-24/27, Mountain View, CA, USA	44

### 3 Approval of minutes of meetings 46 and 47

Input document:

[2903](#) Draft minutes – SC 2/WG 2 meeting 46 – Xiamen, China; Uma/Ksar; 2005-08-22

[2953](#) Draft minutes – SC 2/WG 2 meeting 47 – Sophia Antipolis, France; Uma/Ksar; 2006-02-16

Dr. Umamaheswaran presented briefly the meeting minutes from the previous two meetings. No feedback had been received from anyone on these two documents. Delegates were requested to communicate any editorial errors off line to the recording secretary.

Meeting 46 minutes were adopted without any changes.

Meeting 47 minutes were adopted with the following change (communicated by Prof. Kyongsok Kim off line):

The new action item M47-2c (in section 14.3) on the convener is modified to include the specific document N[2918](#) that was to be communicated to DPRK. The first sentence in action item is modified as:

"Add the topic on Hangul clarification to next meeting's agenda and arrange to get document N2918 and the following resolution sent to DPRK for their review and feedback ...."

### 4 Review action items

Input document:

[2953](#) Minutes – meeting 47 ; Uma/Ksar; 2006-02-16

Dr. Umamaheswaran reviewed the outstanding and new action items from section 14 of M47 meeting minutes in document N[2953](#). Of the 4 outstanding action items all were completed. Of the 49 new action items from meeting M48, 47 were completed and 2 are carried over.

#### 4.1 Outstanding action items from meeting 43, 2003-12-09/12, Tokyo, Japan

Item	Assigned to / action (Reference resolutions in document N2554, and unconfirmed minutes in document N2553 for meeting 43, with any corrections noted in section 3 of document N2653 from meeting 44)	Status
<b>AI-43-9</b>	<b>The Irish national body (Mr. Michael Everson)</b>	
<b>d</b>	To update the roadmap to provide some guidelines for differentiating Ideographic Strokes from Ideographs addressing concerns expressed in document N2526. M44, M45, M46, and M47 – in progress.	Reassigned to US national body (Ken Whistler); Ken reported this item is completed; text will be added to roadmap.

#### 4.2 Outstanding action items from meeting 46, 2005-01-24/28, Xiamen, China

Item	Assigned to / action (Reference resolutions in document N2904, and unconfirmed minutes in document N2903 for meeting 46 - with any corrections noted in section 3 of document N3103 from meeting 48).	Status
<b>AI-46-6</b>	<b>IRG Rapporteur (Dr. Lu Qin)</b>	
	To take note of and act upon the following items.	
<b>c.</b>	To review and feedback on document N2864 - Proposal to add a block of CJK Basic Strokes to the UCS; Thomas Bishop and Richard Cook – individual contribution; 2004-10-25, in view of the text accepted for FDAM1. M47 – in progress.	Completed; covered by document N3063.
<b>AI-46-8</b>	<b>Chinese national body (Messrs, Chen Zhuang, Woushoeur Silamu)</b>	
	M46.30 (Additional characters for Uighur, Kazakh and Kirghiz): With reference to proposed additions to the Arabic script in document N2897 to support Uighur, Kazakh and Kirghiz, WG 2 invites China to take into account the ad hoc discussions during meeting 46, and provide the requested clarification on rendering rules. M47 – in progress.	Completed; covered by documents N2931 and N2992
<b>AI-46-10</b>	<b>All national bodies and liaison organizations</b>	
	To take note of and act upon the following items.	
<b>e.</b>	To take note to restrict their ballot comments to the content of the document under ballot. All proposals for new characters or new material for the standard should be made as independent contributions outside the ballot comments. The project editor has the prerogative of ruling such proposals to be 'out of order, as not related to the text under ballot' and ignore them completely. (To remind national bodies one more time).	Noted; SC 2 also sent a reminder.

#### 4.3 New action items from meeting 47, 2005-09-12/15, Sophia Antipolis, France

Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47)	Status
<b>AI-47-1</b>	<b>Meeting Secretary - Dr. V.S. UMaaheswaran</b>	
<b>a.</b>	To finalize the document N2954 containing the adopted meeting resolutions and send it to the convener as soon as possible.	Completed; see document N2954.
<b>b.</b>	To finalize the document N2953 containing the unconfirmed meeting minutes and send it to the convener as soon as possible.	Completed; see document N2953.
<b>c.</b>	To alert and ensure WG 2 adopts a resolution that was missed at M47 for 12 characters accepted for Amendment 3: <i>M48.x (Latin phonetic and orthographic characters):</i> With reference to documents N2945 on additional Latin phonetic and orthographic characters, WG 2 accepts the following seven (7) Latin characters and five (5) Modifier Tone Letters, with their glyphs as shown in document N2945, for inclusion in Amendment 3 to the standard: 2C6D LATIN CAPITAL LETTER ALPHA 2C6E LATIN CAPITAL LETTER M WITH HOOK 2C6F LATIN LETTER TRESILLO 2C70 LATIN LETTER CUATRILLO 2C71 LATIN SMALL LETTER V WITH RIGHT HOOK	Completed; on the agenda.

Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47)	Status
	2C72 LATIN CAPITAL LETTER W WITH HOOK 2C73 LATIN SMALL LETTER W WITH HOOK in the Latin Extended-C block, and A71B MODIFIER LETTER RAISED UP ARROW A71C MODIFIER LETTER RAISED DOWN ARROW A71D MODIFIER LETTER RAISED EXCLAMATION MARK A71E MODIFIER LETTER RAISED INVERTED EXCLAMATION MARK A71F MODIFIER LETTER LOW INVERTED EXCLAMATION MARK in the Modifier Tone Letters block.	
<b>AI-47-2</b>	<b>Convener - Mr. Mike Ksar</b>	
a.	<i>M47.37 (Roadmap snapshot):</i> WG 2 instructs its convener to post an updated snapshot of the roadmaps (document N2986) as soon as it is ready to the WG 2 web site.	Completed; document N2986 is posted to WG 2 site.
b.	To communicate to SC 2 secretariat on <i>M47.31 (10646 free availability):</i> WG 2 requests SC 2 to request to JTC 1 to make ISO/IEC 10646: 2003 and its Amendment s freely available on the web.	Completed via resolution document.
c.	Add the topic on Hangul clarification to next meeting's agenda and arrange to get document N2918 and the following resolution sent to DPRK for their review and feedback: <i>M47.28 (Hangul clarification):</i> With reference to documents N2994 and N2996, WG 2 recognizes there is a need for some clarification regarding Hangul syllables and invites interested experts to participate in the discussion. WG 2 also instructs its editor to draft appropriate clarification text to address the different items in these documents for further discussion and adding to the agenda for the next meeting.	Completed; on the agenda.
d.	To add the following carried forward scripts to next meeting agenda: <ul style="list-style-type: none"> <li>a. Manichaeian script – document N2544</li> <li>b. Avestan and Pahlavi – document N2556</li> <li>c. Dictionary Symbols – document N2655</li> <li>d. Myanmar Additions – documents N2768, N2827, N2966</li> <li>e. Invisible Letter – document N2822</li> <li>f. Palestinian Pointing characters - document N2838</li> <li>g. Babylonian Pointing characters – document N2759</li> <li>h. Bantu Phonetic Click characters – document N2790</li> <li>i. Samaritan Pointing - 12 Hebrew characters – document N2758</li> <li>j. Dominoe Tiles and other Game symbols – document N2760</li> <li>k. Four Tibetan characters for Balti -- document N2985</li> <li>l. Mahjong Symbols -- document N2975</li> <li>m. Medievalist characters -- document N2957</li> <li>n. Carian, Lycian and Lydian scripts -- documents N2938, N2939 and N2940</li> <li>o. Unconfirmed minutes of Meeting 46 -- document N2903</li> </ul>	Documents N2903 (M46 minutes), document 3043 (on Myanmar topic covers item d), N2985 (Tibetan for Balti) are on agenda; document N2975R of 1006-04-19 on Mahjong Symbols ????. Rest carried forward.
<b>AI-47-3</b>	<b>Editor of ISO/IEC 10646: Mr. Michel Suignard with assistance from contributing editors</b>	
	<i>To prepare the appropriate Amendment texts, sub-division proposals, collection of editorial text for the next edition, corrigendum text, or entries in collections of characters for future coding, with assistance from other identified parties, in accordance with the following:</i>	Completed; see documents SC 2/N3832 and SC 2/N3823 on FDAM2 and documents SC 2/N3816 (WG 2 N2999) and SC 2/N3817 on PDAM3.
a.	<i>M47.5 (Glottal Stop, Claudian and other characters):</i> With reference to documents N2942, N2960, and N2962, WG 2 accepts the following reassignment of fourteen (14) characters and addition of eleven (11) new characters in Amendment 2 to the standard: <ul style="list-style-type: none"> <li>a. Move the characters currently encoded in positions from 0242 to 024E down</li> </ul>	



Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47)	Status
	<p>by one code position to 0243 to 024F, and move the character currently encoded at 024F to 2C74.</p> <ul style="list-style-type: none"> <li>b. 0242 - LATIN SMALL LETTER GLOTTAL STOP; with its glyph from document N2962 in the Latin Extended block; (at the code position 0242 vacated as a result of a. above)</li> <li>c. 037B - GREEK SMALL REVERSED LUNATE SIGMA SYMBOL; with its glyph from document N2991 in the Greek and Coptic block</li> <li>d. 037C - GREEK SMALL DOTTED LUNATE SIGMA SYMBOL; with its glyph from document N2991 in the Greek and Coptic block</li> <li>e. 037D - GREEK SMALL REVERSED DOTTED LUNATE SIGMA SYMBOL; with its glyph from document N2991 in the Greek and Coptic block.</li> <li>f. 04CF - CYRILLIC SMALL LETTER PALOCHKA; with its glyph from document N2991 in the Cyrillic block</li> <li>g. 214E - TURNED SMALL F; with its glyph from document N2962 in the Letterlike Symbols block</li> <li>h. 2184 - LATIN SMALL LETTER REVERSED C; with its glyph from document N2962 in the Number Forms block</li> <li>i. 2C75 - LATIN CAPITAL LETTER HALF H; with its glyph from document N2962 in the Latin Extended-C block</li> <li>j. 2C76 - LATIN SMALL LETTER HALF H; with its glyph from document N2962 in the Latin Extended-C block</li> <li>k. 2C65 - LATIN SMALL LETTER A WITH STROKE; with its glyph from document N2991 in the Latin Extended-C block</li> <li>l. 2C66 - LATIN SMALL LETTER T WITH DIAGONAL STROKE; with its glyph from document N2991 in the Latin Extended-C block</li> </ul>	
<b>b.</b>	<p><i>M47.6 (Additional Cyrillic characters):</i>  With reference to document N2933, WG 2 accepts the following ten (10) additional Cyrillic characters, with their glyphs as shown in document N2933 page 3, and modified during resolution of Irish ballot comment T.1 in document 2959, for inclusion in Amendment 2 to the standard:</p> <ul style="list-style-type: none"> <li>04FA CYRILLIC CAPITAL LETTER GHE WITH STROKE AND HOOK</li> <li>04FB CYRILLIC SMALL LETTER GHE WITH STROKE AND HOOK</li> <li>04FC CYRILLIC CAPITAL LETTER HA WITH HOOK</li> <li>04FD CYRILLIC SMALL LETTER HA WITH HOOK</li> <li>04FE CYRILLIC CAPITAL LETTER HA WITH STROKE</li> <li>04FF CYRILLIC SMALL LETTER HA WITH STROKE</li> </ul> <p>in the Cyrillic block, and</p> <ul style="list-style-type: none"> <li>0510 CYRILLIC CAPITAL LETTER REVERSED ZE,</li> <li>0511 CYRILLIC SMALL LETTER REVERSED ZE,</li> <li>0512 CYRILLIC CAPITAL LETTER EL WITH HOOK,</li> <li>0513 CYRILLIC SMALL LETTER EL WITH HOOK</li> </ul> <p>in the Cyrillic Supplement block.</p>	
<b>c.</b>	<p><i>M47.7 (Additional Latin characters in support of Uighur &amp; Kazakh):</i> With reference to documents N2931 and N2992, on additional Latin characters in support of Uighur, WG 2 accepts the following six (6) Latin characters, with their glyphs as shown in document N2931 on page 3, for inclusion in Amendment 2 to the standard:</p> <ul style="list-style-type: none"> <li>2C67 - LATIN CAPITAL LETTER H WITH DESCENDER</li> <li>2C68 - LATIN SMALL LETTER H WITH DESCENDER</li> <li>2C69 - LATIN CAPITAL LETTER K WITH DESCENDER</li> <li>2C6A - LATIN SMALL LETTER K WITH DESCENDER</li> <li>2C6B - LATIN CAPITAL LETTER Z WITH DESCENDER, and</li> <li>2C6C - LATIN SMALL LETTER Z WITH DESCENDER</li> </ul> <p>in the Latin Extended-C block</p>	
<b>d.</b>	<p><i>M47.8 (Dotted Square):</i> With reference to document N2943, WG 2 accepts for inclusion in Amendment 2:</p> <ul style="list-style-type: none"> <li>2B1A - DOTTED SQUARE; with its glyph as shown in document N2943, in the Miscellaneous Symbols and Arrows block.</li> </ul>	
<b>e.</b>	<p><i>M47.9 (Bottom Tortoise Shell Bracket):</i>  With reference to document N2842 and the Irish ballot comment T.2 in document N2959 regarding Bottom Tortoise Shell Bracket, WG 2 accepts for inclusion in Amendment 2:</p> <ul style="list-style-type: none"> <li>23E1 BOTTOM TORTOISE SHELL BRACKET, with glyph shown in document N2842, and move the character ELECTRICAL INTERSECTION currently encoded at 23E1 to new code position 23E7 in the Miscellaneous Technical block.</li> </ul>	



Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47)	Status
f.	<p><i>M47.10 (Phags-Pa):</i> With reference to the Phags-Pa glyph corrections in document N2972 and N2979, WG 2 accepts the new glyphs proposed in document N2991 for A843, A844, A845, A852, A856, A857, A859, A863, A864, A867, A868, A870, and A871 for Amendment 2. WG 2 also accepts (with reference to editor's note E.2 on the last page of document N2980) to remove the erroneous word 'SMALL' in the Description of variant appearance column (on pages 1 and 2 of FPDAM-2 text) in the second, third, fourth and fifth entries in the table describing Mongolian variation sequences.</p>	
g.	<p><i>M47.11 (Uralicist characters):</i>  With reference to document N2989, WG 2 accepts six (6) Uralicist phonetic characters for encoding in the standard:  1DFE - COMBINING LEFT ARROWHEAD ABOVE  1DFF - COMBINING RIGHT ARROWHEAD AND DOWN ARROWHEAD BELOW  in the Combining Diacritical Marks Supplement block,  27CA - VERTICAL BAR WITH HORIZONTAL STROKE  in the Miscellaneous Mathematical Symbols-A block, and  2C77 - LATIN SMALL LETTER TAILLESS PHI  in the Latin Extended-C block, and  A720 - MODIFIER LETTER STRESS AND HIGH TONE  A721 - MODIFIER LETTER STRESS AND LOW TONE  in a new Latin Extended-D block from A720 to A7FF  In view of the technical completeness of the proposal and the urgency expressed by the user community for the need of these characters to complete the Uralic Phonetic Alphabet set of characters, WG 2 further resolves to include these in Amendment 2 to the standard.</p>	
h.	<p><i>M47.12 (Cuneiform script):</i> With reference to changes requested on the Cuneiform script in the ballot comments from several national bodies in document N2959, WG 2 accepts to replace tables 199-205 Rows 20-24 Cuneiform in Amendment 2, with new code table and names list (including name changes and character reassignments) based on Irish national body comments in document N2959. The final list of changes is in document N2990 (pages 7 and 8) and the revised charts are in document N2991 (pages 456 to 468).</p>	
i.	<p><i>M47.13 (Redundant line collection 340 definition):</i>  WG 2 accepts to remove the redundant entry D4 C1 in definition of collection 340 and include this correction in Amendment 2.  <i>Note: '340' was erroneously recorded as '370' in the approval resolutions document N2954</i></p>	
j.	<p><i>M47.14 (Glyph correction for Malayalam digit zero):</i> WG 2 accepts the proposed corrected glyph for U+0D66 MALAYALAM DIGIT ZERO similar to that in document N2971 and to include the corrected glyph in Amendment 2.</p>	
k.	<p><i>M47.15 (Defect in names of Tai Xuan Jing symbols):</i> With reference to technical defect in names of Tai Xuan Jing symbols reported in document N2988, WG 2 resolves to add an informative annotation and additional text in Annex P, based on input in document N2998 for inclusion in Amendment 2.</p>	
l.	<p><i>M47.16 (Miscellaneous glyph defects):</i> WG 2 accepts the proposed corrected glyphs for 33AC, 06DF, 06E0, 06E1, 17D2 and 10A3F (taking note of the new convention of using dashed square boxes to indicate invisible characters, from document N2956, and 1D09F and 1D09C (from US T.7 resolution in document N2990) for inclusion in Amendment 2.</p>	
m.	<p><i>M47.17 (N'Ko):</i> WG 2 accepts new names for the following four N'Ko characters for inclusion in Amendment 2 based on Irish comment T.1 in document N2959:  07E8 NKO LETTER JONA JA  07E9 NKO LETTER JONA CHA  07EA NKO LETTER JONA RA  07F6 NKO SYMBOL OO DENNEN</p>	
n.	<p><i>M47.18 (Progression of Amendment 2):</i>  WG 2 instructs its editor to prepare the FDAM text based on the disposition of comments in document N2990, the updated code charts in document N2991, and other resolutions from this meeting for additional text for inclusion in Amendment 2 to ISO/IEC 10646: 2003, and forward to SC 2 secretariat for FDAM processing, with unchanged schedule.  (Note: the total number of characters added in FDAM2 has changed to 1365.)  (Items AI-47-3 a through n above are all for Amd 2.)</p>	
o.	<p><i>M47.19 (Inverted Interrobang character):</i></p>	

Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47)	Status
	With reference to document N2935 on Inverted Interrobang character, WG 2 accepts to encode in a future Amendment to the standard: 2E18 - INVERTED INTERROBANG; with its glyph as shown in document N2935 in the Supplemental Punctuation block.	
p.	<i>M47.20 (Musical symbols):</i> With reference to document N2983, WG 2 resolves to: a. Add appropriate informative text in Annex P for U+1D13A MUSICAL SYMBOL MULTI REST to indicate that it is used as a rest corresponding in length to a breve note, which is usually called double rest in American usage or breve rest in British usage. b. Add 1D129 MUSICAL SYMBOL MULTIPLE MEASURE REST with glyph as shown in document N2983 in the Musical Symbols block. in a future Amendment to the standard.	
q.	<i>M47.21 (Malayalam numerals):</i> With reference to document N2970, WG 2 accepts the following six (6) Malayalam numerals for inclusion in a future Amendment to the standard: 0D70 MALAYALAM NUMBER TEN 0D71 MALAYALAM NUMBER ONE HUNDRED 0D72 MALAYALAM NUMBER ONE THOUSAND 0D73 MALAYALAM FRACTION ONE QUARTER 0D74 MALAYALAM FRACTION ONE HALF 0D75 MALAYALAM FRACTION THREE QUARTERS in the Malayalam block, with glyphs similar to those in figure 2 of document N2970.	
r.	<i>M47.22 (Sindhi characters):</i> With reference to document N2934, WG 2 accepts the following four (4) Sindhi characters for inclusion in a future Amendment to the standard: 097B - DEVANAGARI LETTER GGA 097C - DEVANAGARI LETTER JJA 097E - DEVANAGARI LETTER DDDA 097F - DEVANAGARI LETTER BBA in the Devanagari block, with glyphs similar to those in figure 2 of document N2934.	
s.	<i>M47.23 (Greek epigraphical characters):</i> With reference to document N2946, WG 2 accepts the following six (6) Greek epigraphical characters for inclusion in a future Amendment to the standard: 0370 - GREEK CAPITAL LETTER HETA 0371 - GREEK SMALL LETTER HETA 0372 - GREEK CAPITAL LETTER ARCHAIC SAMPI 0373 - GREEK SMALL LETTER ARCHAIC SAMPI 0376 - GREEK CAPITAL LETTER PAMPHYLIAN DIGAMMA 0377 - GREEK SMALL LETTER PAMPHYLIAN DIGAMMA in the Greek and Coptic block, with glyphs based on those in document N2946.	
t.	<i>M47.24 (Lepcha script):</i> WG 2 resolves to encode the Lepcha script proposed in document N2947 in a new block 1C00 – 1C4F named ‘Lepcha’, in a future Amendment to the standard, and populate it with seventy four (74) characters (some of which are combining) with the names, glyphs and code position assignments as shown on pages 17 and 18 in document N2947.	
u.	<i>M47.25 (Vai script):</i> WG 2 resolves to encode the Vai script proposed in document N2948 in a new block A500 – A61F named ‘Vai’, in a future Amendment to the standard, and populate it with two hundred and eighty four (284) characters with the names, glyphs and code position assignments as shown on pages 22 through 27 in document N2948.	
v.	<i>M47.26 (Saurashtra script):</i> WG 2 resolves to encode the Saurashtra script proposed in document N2969 in a new block A880 – A8DF named ‘Saurashtra’, in a future Amendment to the standard, and populate it with eighty one (81) characters (some of which are combining) with the names, glyphs and code position assignments (A880 to A8C4 and A8CE to A8D9) as shown on pages 11 and 12 in document N2969.	
w.	<i>M47.27 (OI Chiki script):</i> WG 2 resolves to encode the OI Chiki script proposed in document N2984 in a new block 1C50–1C7F named ‘OI Chiki’, in a future Amendment to the standard, and populate it with forty eight (48) characters with the names, glyphs and code position assignments as shown on pages 11 and 12 in document N2984.	
x.	<i>M47.29 (Amendment 3 – subdivision and PDAM text):</i> WG 2 instructs its editor to prepare a project sub division proposal and PDAM text based on the various resolutions from this meeting accepting characters for inclusion in a future Amendment	

Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47)	Status
	to ISO/IEC 10646: 2003, and forward them to the SC 2 secretariat for simultaneous ballot of sub division approval and PDAM registration (see documents N2993 and N2999). The proposed completion dates for the progression of this work item are: PDAM 2006-02-15, FPDAM 2006-09-15, and FDAM 2007-02. (Note: A total of 517 characters were accepted for inclusion in Amendment 3 at this meeting). Items AI-47-o through x are all for Amd 3.	
y.	M47.28 (Hangul clarification): With reference to documents N2994 and N2996, WG 2 recognizes there is a need for some clarification regarding Hangul syllables and invites interested experts to participate in the discussion. WG 2 also instructs its editor to draft appropriate clarification text to address the different items in these documents for further discussion and adding to the agenda for the next meeting.	Completed; see document N3006.
<b>AI-47-4</b>	<b>Ad hoc group on principles and procedures (lead - Dr. V.S. UMaaheswaran)</b>	
	To take note of and act upon the following items.	Completed; see document N3002.
a.	<p>M47.1 (Case Folding stability): WG 2 acknowledges the need for guaranteeing stability of case folding of characters in scripts having two cases, and accepts to adopt the guidelines in the following text in its principles and procedures:</p> <p><i>Stability of Case Folding</i></p> <p>For text containing characters from scripts having two cases (bicameral scripts), case folding is an essential ingredient in case-insensitive comparisons. Such comparisons are widely applied, for example in Internationalized Domain name (IDN) lookup, or for identifier matching in case-insensitive programming and mark-up languages. Because such operations require stable identifiers and dependable comparison results it is important that the standard be able to provide the guarantee of complete stability. For historic reasons, and because there are many characters with lowercase forms but no uppercase forms, the case folding is typically done by a <i>lowercasing</i> operation, and this also matches the definition used by the Unicode Standard for <i>case folding</i>.</p> <p>In order to guarantee the case-folding stability, WG 2 adopts the following principle when evaluating proposals for encoding characters for bicameral scripts:</p> <p>"Subsequent to the publication of Amendment 2 of ISO/IEC 10646: 2003, if a character already has an uppercase form in the standard but no lowercase form, its corresponding lowercase form can not be added to the standard; the only way a lowercase form can be added, if proved to be absolutely essential, is by entertaining a new pair of uppercase and lowercase forms for encoding. If only a lowercase form exists and an uppercase form is deemed to be needed it can be added without affecting the stability of case folding."</p>	
b.	M47.2 (Combining diacritic marks): With reference to documents N2976 and N2978, WG 2 acknowledges the need for additional guidelines when considering generic diacritical marks that may be used across different scripts, and accepts the text proposed in document N2978 as additional guidelines for inclusion in its principles and procedures.	
c.	M47.3 (Adding to FDAM): With reference to document N2997, WG 2 acknowledges the need for additional guidelines on handling requests with some urgency in providing solutions, and accepts the text proposed in document N2997, modified per discussions at M47, for inclusion in its principles and procedures. Note: the second reference to document N2997 was erroneously recorded as N2987 in the approved resolution document N2954.	
d.	M47.4 (Disunification, Stability of Identifiers): With reference to document N2987, WG 2 accepts the text proposed in document N2987, modified per discussions at M47, for inclusion in its principles and procedures.	
e.	M47.36 (Principles and procedures): WG 2 accepts the draft updates to its principles and procedures as presented in document N2952 and instructs its convener to post a revised document N3002 based on discussions at M47 and the proposal summary form to the WG 2 web site.	
<b>AI-47-5</b>	<b>IRG Rapporteur (Dr. Lu Qin)</b>	
	To take note of and act upon the following items.	
a.	M47.32 (CJK Ext C2): WG 2 takes note of item 3 in document N2968 and authorizes IRG to start its deliberations regarding CJK Extension C2.	Noted.

Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47)	Status
b.	With reference to discussion in meeting 47, regarding IICORE and safe characters for security, IRG is requested to review and feedback on UTS 36 for safe characters and to RFC 3743.	In progress.
<b>AI-47-6</b>	<b>TCA (Mr. Shih-ShyengTseng)</b>	
a.	To prepare a revised proposal on Mahjong symbols based addressing the concerns and feedback received at meeting M47 on document N2975.	Completed; see revised N2975R.
<b>AI-47-7</b>	<b>Chinese national body (Messrs, Cheng Zhuang)</b>	
a.	To take note of resolution <i>M47.30 (Todo Mongolian punctuation marks)</i> : With reference to document N2963 on three punctuation marks for Todo Mongolian, WG 2 rejects the request since these characters can be unified with already coded characters in the standard as follows: TODO COMMA ONE with 002C - COMMA TODO COMMA TWO with 3001 - IDEOGRAPHIC COMMA TODO FULL STOP with 3002 - IDEOGRAPHIC FULL STOP	Noted.
<b>AI-47-8</b>	<b>Ireland (Mr. Michael Everson)</b>	
a.	To request Ms. Deborah Anderson to send supporting documentation for WG 2 records regarding UNESCO project about N'Ko script; and to get the authors of N'Ko proposal to provide more evidence of simultaneous use of old and new forms included in the script.	Completed; convener to post the document.
b.	To communicate to Myanmar on the status of their documents N2768, N2827, N2966 in WG 2.	Completed; see document N3043.
c.	To work with the author Mr. Nick Nicholas to provide supporting examples for Greek epigraphical characters proposed in document N2946, and accepted per resolution M47.23.	In progress.
<b>AI-47-9</b>	<b>All national bodies and liaison organizations</b>	
	<i>To take note of and act upon the following items.</i>	
a.	<i>M47.33 (Future meetings)</i> : WG 2 meetings: Meeting 48 – 2006-04-24/28, Mountain View, US; along with SC 2 plenary Meeting 49 – 2006-09-25/28, Tokyo, Japan Meeting 50 – Spring 2007, Europe (seeking Host) IRG meeting: IRG #25 - 2005-11-28/12-02, Berkeley, CA, USA (host: Unicode Consortium) IRG #26 - 2006-06-12/16, Hue, Vietnam IRG #27 - 2006-11-27/12-01 (to be confirmed), Taipei, Taiwan (host: TCA) IRG #28 - 2007 May / June, Hangzhou, China (place to be confirmed)	Noted.
b.	<i>M47.28 (Hangul clarification)</i> : With reference to documents N2994 and N2996, WG 2 recognizes there is a need for some clarification regarding Hangul syllables and invites interested experts to participate in the discussion. WG 2 also instructs its editor to draft appropriate clarification text to address the different items in these documents for further discussion and adding to the agenda for the next meeting.	Completed; see documents N3006 and N3045.
c.	To review and feedback to the editor on document N2937: Working Draft ISO/IEC 10646: 2003 plus Amendment 1 merged and combined; Michel Suignard; 2005-04-11.	Noted; no feedback.
d.	To take note of important changes in document N3002 and new format for proposal summary form, prepared according to resolution <i>M47.36 (Principles and procedures)</i> : WG 2 accepts the draft updates to its principles and procedures as presented in document N2952 and instructs its convener to post a revised document N3002 based on discussions at M47 and the proposal summary form to the WG 2 web site.	Noted.
e.	In addition to the proposals mentioned in the items above, the following proposals have been carried forward from earlier meetings. All national bodies and liaison organizations are invited to review and comment on them. a. Manichaeian script – document N2544 b. Avestan and Pahlavi – document N2556 c. Myanmar Additions – documents N2768, N2827, N2966 d. Dictionary Symbols – document N2655 e. Invisible Letter – document N2822 f. Palestinian Pointing characters - document N2838 g. Babylonian Pointing characters – document N2759 h. Bantu Phonetic Click characters – document N2790 i. Samaritan Pointing - 12 Hebrew characters – document N2758 j. Dominoe Tiles and other Game symbols – N2760 k. Four Tibetan characters for Balti -- document N2985	Myanmar (item c) and Tibetan for Balti (item k) have new contributions and are on the agenda; document N2975R of 1006-04-19 on Mahjong Symbols; items m and

Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47)	Status
	<ul style="list-style-type: none"> <li>l. Mahjong Symbols -- document N2975</li> <li>m. Medievalist characters -- document N2957</li> <li>n. Carian, Lycian and Lydian scripts -- documents N2938, N2939 and N2940</li> </ul>	n.

## 5 JTC 1 and ITTF matters

### 5.1 ITU SG17 Requirements for future editions

Input documents:

- [3013](#) SG 17 requirements for future editions of ISO/IEC 10646 - Liaison Statement; ITU-T Study Group 17 (Geneva, 5-14 October 2005); 2005-10-28
- [3026](#) Liaison Statement from ITU-T SG 17 regarding Future editions of ISO/IEC 10646; ITU-T SG-17 SC02n3820; 2005-11-01

Mr. Mike Ksar: SC 2 has received a note from ITU on ITU making ISO/IEC 10646 into an ITU recommendation. They have demanded some textual changes to ISO/IEC 10646. I would like to discuss this topic. There is no one from ITU-T to present their side.

Discussion:

- a. Mr. Michel Suignard: Their request is to have common text. They are also asking for names for ASCII controls. We addressed this under agenda item 9.2 on page 35. It is pretty easy for us to add normative definitions for UTF16BE etc. in the standard, or reference to the Unicode standard. For items c, d and e in their request, we can have a normative reference to the character properties to the database.
- b. Dr. Asmus Freytag: Items c, d and e to some extent refer to the identity of the characters. ISO/IEC 10646 could use normative references to the Unicode standard for these. However, we should not reference the Unicode database as a unit because there are lots of other things in it. The General Category property would address items c and d. Also a reference to the *default* case folding in Unicode would satisfy the user community. As to item f, the Unicode Consortium would defend its rights to the Unicode Trademark fiercely.
- c. Mr. Erkki Kolehmainen: I am in favour of such references. But these should not be normative. Informative references to sources of information should be sufficient.
- d. Mr. Michel Suignard: We already have reference to the BiDi algorithm, which in turn has to use character properties anyway. It is the same case with Normalization. We could add some minor text to satisfy the request, and not cause too much burden for national bodies.

Disposition: The consensus of the discussion is captured in the relevant resolutions M48.34 and M46.38 below. See also relevant resolution M48.33 in section 9.2 on page 35.

*Relevant resolutions:*

*M48.34 (Additions of UTF-16BE and other forms):* *Unanimous*  
 WG 2 accepts the request from ITU-T SG17 in document N3013, item b, to add explicit text for UTF-16BE and other similar forms, and instructs its editor to prepare appropriate text for inclusion in the standard.

*M48.38 (Request from ITU-T SG17):* *Unanimous*  
 WG 2 has reviewed the document N3013 from ITU-T SG17 and has the following responses for SC 2 to consider:

1. ITU-T should be discouraged from the proposed common-text approach and pursue referencing ISO/IEC 10646, similar to what ITU-T has done with other SC 2 standards in the past; for example, a normative reference can be added to recommendation such as T.51.
2. As to their requests a through f:
  - item a: ISO/IEC 10646 already references ISO/IEC 6429 in a normative way. Further new text will be added to the standard to facilitate ease of reference of the control characters from ISO/IEC 10646.
  - item b: New text will be added to clarify explicitly the definitions of terms such as UTF-16BE in the standard.
  - items c, d and e: We need further clarification on the nature of the request from ITU-T SG17. WG 2 requests SC 2 to invite ITU-T representatives to participate in the next WG 2 meeting. The project editor will review these items and draft appropriate text for discussion at the next WG 2 meeting.
  - item f: WG 2 agrees with ITU-T SG17 that there are trademark issues and we are not in favour of changing the expansion of the term UCS.
3. WG 2 nominates its convener Mr. Mike Ksar to be the liaison representative to ITU-T and requests SC 2 to approve this nomination, and to prepare a liaison statement communicating SC 2's response in more detail, along with a reference to document N3057 regarding the Unicode Consortium's status as an ARO.



## 5.2 Vocabulary work group

Input document:

[02N3853](#) Report to ISO/IEC JTC 1/SC 2 on the First Meeting of the JTC 1 Ad Hoc Group on IT Vocabulary on 1-3 March 2006; Mr. Erkki I. Kolehmainen, SC 2 Representative; 2006-04-18

Mr. Erkki Kolehmainen: I attended the ad hoc representing SC2. My report is in SC 2 document [02N3853](#), The meeting was held at the request of JTC 1 hosted by France and chaired by Mr. Joseph Coté of Canada. IT Vocabulary document ISO 2382 is totally out of date. JTC 1/SC 1 has been disbanded a few years ago. The ad hoc was in the direction of updating this standard. The primary expertise does not exist any more in JTC 1. SC 2 has never been directly part of any of the SC 1 standards. SC 2 vocabulary can be included in the standard only if we have a work item in SC 2 to provide the input. It is also highly questionable as to how up to date such an updated standard could be. It could give the wrong impression that JTC 1 has been keeping these up to date when they come up with the next update. The Canadian Translation Bureau has agreed to use their *Termium* database as the tool to keep this work up to date. SC 2 could add their vocabulary to that database. Their next meeting will be this fall before the next JTC 1 plenary. I recommend someone be assigned the task of monitoring this effort and report back to SC 2 and WG 2 as the project evolves on its usefulness / value for us.

Discussion:

- a. Dr. Ken Whistler: What is the value of the vocabulary standard? What happens if it is out of date? Should this committee be worried about lack of SC 2 related terms in that vocabulary standard?
- b. Mr. Alain LaBonté: It is useful for translators and terminologists. I think it is in the interest of SC 2 that whatever definitions we have in our standards are known outside of the users of our standards. The idea is for each SC to input the terminology and associated meanings to the rest of the world. For example, 'repertoire' in the context of SC 2 is not the same as in general usage.
- c. Mr. Mike Ksar: We can go to SC 2 plenary and make some recommendations. It looks like WG 2 does not have any direct interest in this subject.
- d. Mr. Erkki Kolehmainen: I am recommending monitoring the activity and keep SC 2 and WG 2 informed.

## 5.3 Other JTC 1 documents

Input documents:

[3009](#) SC 2 Business Plan for JTC 1 Plenary - 2005; Kobayashi-san, SC 2 Chair; 2005-10-04  
[3017](#) Notice of Publication: ISO/IEC 10646: 2003/Amd. 1; SC 2 Secretariat n3823; 2005-11-22  
[3018](#) JTC 1 Resolutions - November 2005; SC 2 Secretariat n3824; 2005-12-05

The above documents related to JTC 1 and ITTF work were brought to the attention of WG 2 members. These were not discussed at the meeting. The SC 2 business plan was submitted by SC 2 chair to JTC 1.

## 6 SC 2 matters

### 6.1 SC 2 Program of Work

Input documents:

[3034](#) ISO/IEC 10646:2003/Amd.2:2006 as submitted for balloting; Project Editor - Michel Suignard; 2006-02-01  
[3036](#) SC 2 Request to make ISO/IEC 10646 and ISO/IEC 14651 (including subsequent editions and future amendments) to make freely available; SC2 Secretariat; 2006-02-03

The above documents related to work of SC 2 were brought to the attention of WG2 experts. There was no discussion at the meeting. The request 'to make ISO/IEC 10646 and ISO/IEC 14651 standards freely available' has been sent to ISO and IEC boards.

### 6.2 Ballot results - Amendment 3

Input documents:

[3071](#) Summary of Voting/Table of Replies - PDAM 3; SC2 Secretariat, 02n3852; 2006-04-07  
[3072](#) Summary of Voting on SC 2 N 3816: Project Subdivision Proposal for ISO/IEC 10646: 2003/Amendment 3; SC2 Secretariat, 02n3851; 2006-04-07

The Amendment 3 subdivision proposal was approved with 14 national bodies accepting it, and 4 abstaining; 12 had not voted. Of these 7 national bodies indicated they will participate in the development. The PDAM 3 text was approved with 11 national bodies approving it (3 with

comments), 1 disapproving it, 6 abstaining; 12 had not voted. See disposition of ballot comments under item 10.1 on page 38.

## **7 WG 2 matters / housekeeping**

### **7.1 Updates to Roadmap per N2526 - Action item 43-9**

Input document:

[3005](#) Snapshot of Roadmaps; Uma; 2006-04-12

The following text has been added to the [SIP roadmap](#) on the Unicode site resulting from action item AI 43-9 to add some clarification text in response to concerns expressed in document [N2526](#); this addition was done before close of this meeting.

"The SIP is intended for encoding of additional repertoire of Unified CJK ideographs or Compatibility CJK ideographs. Other CJK related characters, including strokes, radicals, and punctuation are encoded on the BMP."

Document [N3005](#) is the updated snapshot of roadmaps prior to this meeting, and does NOT include the above text.

Action item: Ad hoc on roadmap to incorporate the new text in the next snapshot of roadmaps.

Disposition: Accept the snapshot document. There was no discussion.

*Relevant resolution:*

*M48.43 (Roadmap snapshot):*

*Unanimous*

WG 2 instructs its convener to post an updated snapshot of the roadmaps (in document N3005) to the WG 2 web site.

### **7.2 Restrict ballot comments to content under ballot - reminder**

Experts are reminded to take note and abide by the message in action item [AI-46-10](#) on page 5.

### **7.3 Missing resolution from M47 on Latin phonetic & orthographic characters**

The secretary had missed entering the following in the resolutions adopted from meeting M47 - reported in action item AI 47-1c. The characters are included in Amendment 3 based on discussion at the meeting M47. This should be added as a resolution at meeting M48 for completeness of resolutions governing Amendment 3.

Disposition: Accepted.

*Relevant resolution:*

*M48.1 (Latin phonetic and orthographic characters from M47):*

*Unanimous*

With reference to documents N2945 on additional Latin phonetic and orthographic characters, WG 2 accepts the following seven (7) Latin characters and five (5) Modifier Tone Letters, with their glyphs that are similar to those shown in document N2945, for inclusion in Amendment 3 to the standard:

2C6D LATIN CAPITAL LETTER ALPHA

2C6E LATIN CAPITAL LETTER M WITH HOOK

2C6F LATIN LETTER TRESILLO

2C70 LATIN LETTER CUATRILLO

(Note: the above two characters have been subsequently replaced in M48; see resolution M48.2)

2C71 LATIN SMALL LETTER V WITH RIGHT HOOK

2C72 LATIN CAPITAL LETTER W WITH HOOK

2C73 LATIN SMALL LETTER W WITH HOOK

in the Latin Extended-C block, and

A71B MODIFIER LETTER RAISED UP ARROW

A71C MODIFIER LETTER RAISED DOWN ARROW

A71D MODIFIER LETTER RAISED EXCLAMATION MARK

A71E MODIFIER LETTER RAISED INVERTED EXCLAMATION MARK

A71F MODIFIER LETTER LOW INVERTED EXCLAMATION MARK

in the Modifier Tone Letters block.



## 7.4 Update to Principles & Procedures

Input document:

[3002](#) Principles and Procedures for Allocation of New Characters and Scripts and handling of Defect Reports on Character Names (Replaces N2952, N2902, N2652R, N2352R, N 2002 and N1876); Uma - Ad hoc group on Principles & Procedures; 2005-10-05

Document [N3002](#) and associated proposal summary forms with revised formats (no changes to the content) have been posted to the WG 2 web site.

Action item: All national bodies and liaison organizations are to take note. Request SC2 to send out a notification to SC2 member bodies.

## 7.5 WG 2 report to SC 2 plenary

Input document:

[3056](#) WG2 report to SC2 plenary - April 2006; Ksar; 2006-03-06

This document is the convener's report to SC2. It will be updated reflecting the results of this meeting before submission to SC2. There was no discussion.

## 8 New and updated contributions – not part of Amendment 3

Mr. Mike Ksar: I seek your input in prioritizing the various topics under this agenda item.

Discussion:

- a. Dr. Ken Whistler: Some items that are in item 8 do touch upon some comments on Amendment 3. For example: Vai script and Saurashtra items. It may be smoother to go through Amendment 3 disposition of comments first. It would be easier to pick through agenda items that ask for a few characters rather than whole scripts. These could be candidates for inclusion in the current Amendment - for example, the Two bracket characters, four Malayalam characters and others.
- b. Mr. Michael Everson: There is a revised document [N3081R](#) to assist the delegates in reviewing Vai comments. I will give this to the convener prior to its discussion.
- c. Mr. Michel Suignard: As and when we identify the relevant agenda item we can jump to it.
- d. Mr. Mike Ksar: As we go through this list we can go by the above guideline. Our editor thinks that the disposition of comments on Amendment 3 should not take long.

Note: Some of agenda items below in this section are in support of Amendment 3 ballot comments.

### 8.1 Tibetan additions

#### 8.1.1 Four Tibetan characters for Balti

Input documents:

[2985](#) Proposal to add four Tibetan characters for Balti; Michael Everson; 2005-09-05

[3010](#) Comments on Michael Everson's proposal to encode four additional Tibetan characters for Balti (N2985); Andrew West; 2005-10-25

Mr. Michael Everson: The proposed characters in document [N3010](#) are under discussion among experts. Document [N2985](#) is precursor of document [N3010](#). This proposal is for national body feedback.

Action item: National bodies to feedback on document [N3010](#).

#### 8.1.2 One Tibetan astrological character

Input documents:

[3011](#) Proposal to encode one Tibetan astrological character; Andrew West; 2005-10-24

Mr. Michael Everson: Document [N3011](#) proposes a single astrological character. There is enough evidence given now; we did not have that earlier. We should accept this.

Discussion:

- a. Dr. Ken Whistler: US national body has reviewed this and agree to go ahead.
- b. Dr. Asmus Freytag: The UTC has looked at it and came to the conclusion to include it.

Disposition: Accept the astrological character in document [N3011](#). See relevant resolution M48.20 under section 8.1.3 below.

### 8.1.3 Three archaic Tibetan characters

Input documents:

- [3012](#) Proposal to encode three archaic Tibetan characters; Andrew West; 2005-10-24
- [3032](#) Proposal to encode one Tibetan punctuation mark; Andrew West; 2006-01-30
- [3033](#) Proposal to encode two archaic Tibetan punctuation marks; Andrew West; 2006-01-30

Mr. Michael Everson: Document [N3012](#), from Mr. Andrew West proposes three archaic Tibetan characters. Of these three characters one of them is in document [N3032](#). Reasoning for this character is different from the other two; their properties are also different; the new document includes some updates as well. Document [N3033](#) has the remaining two characters.

Discussion:

- a. Dr. Ken Whistler: The US national body has reviewed and discussed these characters at great length and have accepted them.
- b. Dr. Umamaheswaran: Did China input on these Tibetan characters?
- c. Dr. Lu Qin: China had sent an email on this showing disagreement.
- d. Mr. Mike Ksar: There were no written contributions on this from China.
- e. Mr. Michael Everson: There is an opportunity for commenting during the ballot.
- f.

Disposition: Accept the single character from document [N3032](#), and the two characters from document [N3033](#). See relevant resolution M48.20 below.

*Relevant resolution:*

*M48.20 (4 Tibetan characters):*

*Unanimous*

WG 2 accepts the following four Tibetan *characters* to encode in the Tibetan block:

- a. 0FCE TIBETAN SIGN RDEL NAG RDEL DKAR  
with its glyph similar to that shown in document N3011
- b. 0FD2 TIBETAN MARK NYIS TSHEG  
with its glyph similar to that shown in document N3032, and,
- c. 0FD3 TIBETAN MARK INITIAL BRDA RNYING YIG MGO MDUN MA  
0FD4 TIBETAN MARK CLOSING BRDA RNYING YIG MGO SGAB MA  
with their glyphs that are similar to those shown in document N3033,

### 8.2 Kaithi script

Input document:

- [3014](#) Towards an encoding of the Kaithi script in the SMP of the UCS; Michael Everson - Individual Contribution; 2005-11-06

This document is for feedback from national bodies.

Action item: National bodies to review and feedback on document [N3014](#) on Kaithi script..

### 8.3 Malayalam additions

Input document:

- [3015](#) Proposal for addition of 4 characters to Malayalam Block; UTC, R. Chitrajakumar, Rachana Akshara Vedi; 2005-10-24

Dr. Umamaheswaran: The four Malayalam characters proposed in document [N3015](#) are well exemplified. Two are combining marks. The other two are non-combining marks. We should accept these characters -- the names of the last two could be subject to discussion.

Dr. Ken Whistler: The US has no objection to adding these. It is appropriate to add these to Amendment 3. We have a similar character for Kannada.

Dr. Asmus Freytag: Suggested positions can be referenced to 'as in document [N3059](#)' that will contain the updated charts for amendment(s).

Disposition: Accept the four; with annotation (avagraha) for the Praslesham character. See relevant resolution M48.18 below.

*Relevant resolution:*

*M48.18 (4 Malayalam characters):*

*Unanimous*

With reference to document N3015, WG 2 accepts the following four (4) additional Malayalam characters, with their glyphs that are similar to those shown in document N3015 to encode in the Malayalam block:

- 0D44 MALAYALAM VOWEL SIGN VOCALIC RR
- 0D62 MALAYALAM VOWEL SIGN VOCALIC L
- 0D3D MALAYALAM PRASLESHAM (avagraha) ←10000th character
- 0D79 MALAYALAM DATE MARK

## 8.4 Lycian and Lydian scripts

Input document:

[3019](#) Proposal to encode the Lycian and Lydian scripts in the SMP of the UCS; Everson; 2006-01-12

Mr. Michael Everson: Lycian and Lydian scripts are ancient Anatolian scripts. They are related to Greek but have more characters not related well to Greek. It is not appropriate to unify either of these with Greek. Lycian is a Left to Right script. It does not use script-specific punctuation -- uses a two dot punctuation for word division. Samples are shown on pages 2, 3 and 4. Experts on Anatolian languages have reviewed and given feedback. Lydian is also an alphabetic script -- it is Right to Left unlike Lycian. Both these are proposed in SMP. One punctuation mark Lydian Quotation mark is shown in this document. There is a question on its name from the UTC -- the proposal is to call it Lydian Triangular Mark because its use of quotation is not clear.

Ms. Lisa Moore: The UTC has accepted these scripts.

Disposition: Accept Lycian and Lydian scripts for encoding in the SMP. See relevant resolutions M48.6 and M48.7 below.

*Relevant resolutions:*

*M48.6 (Lycian script):*

*Unanimous*

WG 2 resolves to encode the Lycian script proposed in document N3019 in a new block 10280 -1029F in the SMP, named 'Lycian', and populate it with twenty nine (29) characters with the names, glyphs and code position assignments as shown on pages 10 and 11 in document N3019.

*M48.7 (Lydian script):*

*Unanimous*

WG 2 resolves to encode the Right-to-Left Lydian script proposed in document N3019 in a new block 10920 -1093F in the SMP, named 'Lydian', and populate it with twenty seven (27) characters with the names, glyphs and code position assignments as shown on pages 12 and 13 in document N3019.

## 8.5 Carian script

Input document:

[3020](#) Proposal to encode the Carian script in the SMP of the UCS; Everson; 2006-01-12

Mr. Michael Everson: Carian is another one of Anatolian languages. It is a Left to Right script - it is a little more complicated. Uses space and sometimes word dividers. Due to its unusual nature its decipherment process has been a gradual one. There are characters that are known to be variants of others -- some of these have become distinguished over the years of decipherment. For that reason what are variants now were not considered variants in the history of decipherment. The names of the characters have a convention of using -2, 3, etc. in them, similar to Cuneiform script names. An update to the proposal document has been posted.

Dr. Asmus Freytag: The UTC has reviewed and has accepted the proposal.

Disposition: Accept Carian script for encoding in the SMP per document [N3020](#).

*Relevant resolution:*

*M48.8 (Carian script):*

*Unanimous*

WG 2 resolves to encode the Carian script proposed in document N3020 in a new block 102A0 -102DF in the SMP, named 'Carian', and populate it with forty nine (49) characters with the names, glyphs and code position assignments as shown on pages 12 and 13 in document N3020.

## 8.6 Gurmukhi additions

### 8.6.1 Gurmukhi tone mark Udaat

Input document:

[3021](#) Proposed Changes to Gurmukhi; Sukhjinder Sidhu (Punjabi Computing Resource Centre), Unicode Technical Committee; 2005-12-28

Dr. Ken Whistler: This is a request to add one tone mark at 0A51. Rest of the document gives rationale and background discussion. The main issue is that it looks similar to Halant (Virama). The document demonstrates that it is different from Halant, it has a different function and shape etc. It is a missing character - we should support it.

Discussion:

- a. Dr. Umamaheswaran: I support the addition of this character.
- b. Mr. Mike Ksar: Accept.

- c. Mr. Michel Suignard: Which Amendment should we include it in?
- d. Dr. Asmus Freytag: We have one more technical ballot - Amendment 3 should be OK.
- e. Dr. Ken Whistler: we should add this to Amendment 3 -- it already talks about adding a few miscellaneous Indic characters.

Disposition: Accept as proposed for current Amendment. See relevant resolution M48.25 under section 8.6.2 below.

### 8.6.2 Gurmukhi sign Yakash

Input document:

[3073](#) Gurmukhi Sign Yakash; Sukhjinder Sidhu (Punjabi Computing Resource Centre), Unicode Technical Committee; 2006-04-07

Dr. Ken Whistler: This is another character that has turned up based on specialists on Punjabi using Gurmukhi script. This particular sign came out of writing Sanskrit in Gurmukhi script. It is a combining mark that shows up in usage. Evidence of usage is given. We cannot use anything else for it. It is appropriate to encode this. The proposal came from experts.

Disposition: Accept as proposed. See relevant resolution M48.25 below.

*Relevant resolution:*

*M48.25 (2 Gurmukhi characters):*

*Unanimous*

WG 2 accepts the following two (2) Gurmukhi characters for encoding in the Gurmukhi block:

- a. 0A51 GURMUKHI SIGN UDAAT  
with its glyph similar to that shown in document N3021
- b. 0A75 GURMUKHI SIGN YAKASH  
with its glyph similar to that shown in document N3073; this is a combining character,

### 8.7 Sundanese script

Input document:

[3022](#) Proposal for encoding the Sundanese script in the BMP of the UCS; Michael Everson; 2006-01-09

Mr. Michael Everson: Sundanese is another script that is being revived and is being taught in Indonesia. The Prasasti Kawali stone in figure 1 on page 4 of the proposal document was used as reference for the script. There were some archaic characters on the stone that are NOT in the current proposal. This script is being taught in schools. It is a simplified Brahmic script. A modern keyboard is used. European punctuation is used. Being a new script should have full period of review. In the list of names the annotations assist the non-native speakers. The traditional names are not used by current users of the script - annotations will assist. We should keep the annotations. The request is to encode in the BMP.

Disposition: Accept for two technical ballots; as proposed in document [N3022](#). See relevant resolution M48.9 below.

*Relevant resolution:*

*M48.9 (Sundanese script):*

*Unanimous*

WG 2 resolves to encode the Sundanese script proposed in document N3022 in a new block 1B80 -1BBF in the BMP, named 'Sundanese', and populate it with fifty five (55) characters (some of which are combining) with the names, glyphs and code position assignments as shown on pages 7 and 8 in document N3022.

### 8.8 Rejang script

Input document:

[3096](#) Proposal for encoding the Rejang script in the BMP of the UCS; Everson; 2006-04-24

Mr. Michael Everson: Document [N3096](#) replaces the previous preliminary proposal in document [N3023](#). The new document corrects the population information and updates the sorting order information. Rejan script is a Brahmic script - derived from a South Indian script - but is much simplified. It does not have the conjoining behaviour like others. It follows the usual Brahmic order. It has a number of dependant vowel signs and dependant finals. The Virama is used truly as vowel killer -- does not cause any conjoining behaviour - with no other name for it. A special punctuation mark - section mark - is used sometimes; sometimes at the beginning of phrases and sometimes at the end. Examples are on page 4. Otherwise they don't use punctuation marks -- other than spaces and a colon. In the particular font used in this example we have lot of spaces. The font was created in 1964. It is simple, the character set is well known since 1783 and there have been no additions till

the work of 1964. The 'Mark of Pause' is the Virama. The proposal is for 37 characters. It is a contemporary script and is proposed for the BMP.

Disposition: Accept for encoding with two technical ballots. Annotations proposed to identify their traditional names will be removed from the tables in [N3096](#). See relevant resolution M48.10 below.

*Relevant resolution:*

*M48.10 (Rejang script):*

*Unanimous*

WG 2 resolves to encode the Rejang script proposed in document N3096 in a new block A930 -A95F in the BMP, named 'Rejang', in a future Amendment to the standard, and populate it with thirty seven (37) characters (some of which are combining) with the names (without the annotations), glyphs and code position assignments as shown on pages 5 and 6 in document N3096.

## **8.9 Kayah Li script**

Input document:

[3038](#) Proposal for encoding the Kayah Li script in the BMP of the UCS; Everson; 2006-03-09

Mr. Michael Everson: Document [N3038](#) replaces the previous preliminary proposal in document [N3024](#). Kayah Li is a modern script invented in 1962 based on Latin orthographies to write Eastern and Western Kayah Li languages. These are used both in Myanmar and in Thailand. They are not related to Myanmar script. It is an alphabet -- it is not Brahmic. Fairly small, fonts, keyboards etc. are available. It consists of 48 characters - a set of letters, set of vowel signs and set of tone marks. Some of the tone marks look like they could be composed of others. The consensus is not to compose those that look like being composable. There are also ten digits. The script is proposed for inclusion in the BMP.

Discussion:

- a. Mr. Michel Suignard: There is no more to add?
- b. Mr. Michael Everson: The character set has been stable for a long time. Charts were out before the formal proposal has been made. There are two punctuation marks that were unknown to us earlier. This is a first time submission to WG 2.
- c. Mr. Mike Ksar: Being the first time proposal it should not go into current Amendment.
- d. Dr. Asmus Freytag: UTC has looked at this and would like to have at least two technical ballots on this proposal.
- e. Mr. Michael Everson: I concur that this proposal should have at least two technical ballots.

Disposition: Accept the proposal as in document [3038](#) (a revised document has been posted); for two technical ballots. See relevant resolution M48.11 below.

*Relevant resolution:*

*M48.11 (Kayah Li script):*

*Unanimous*

WG 2 resolves to encode the Kayah Li script proposed in document N3038 in a new block A900 -A92F in the BMP, named 'Kayah Li', and populate it with forty eight (48) characters (some of which are combining) with the names, glyphs and code position assignments as shown on pages 7 and 8 in document N3038.

## **8.10 Medievalist characters**

Input documents:

[3027](#) Proposal to add medievalist characters to the UCS; Michael Everson (editor), Peter Baker, António Emiliano, Florian Grammel, Odd Einar Haugen, Diana Luft, Susana Pedro, Gerd Schumacher, Andreas Stötzner; 2006-01-30

[3039](#) Feedback on N3027 Proposal to add Medievalist Characters; Unicode Technical Committee and US National Body; 2006-03-16

[3060](#) Feedback on N3027 - medievalist characters; UK NB - Andrew West; 2006-03-27

[3077](#) Response to UTC/US contribution N3037R, "Feedback on N3027 Proposal to add medievalist characters"; Michael Everson (editor), Peter Baker, António Emiliano, Florian Grammel, Odd Einar Haugen, Diana Luft, Susana Pedro, Gerd Schumacher, Andreas Stötzner; 2006-03-31

Document [N3027](#) is the original proposal. The other documents in the above list are comments and responses on the original.

An ad hoc group of interested experts met and reported back.

Dr. Ken Whistler: The consensus in the ad hoc was to delete 1 character based on the US comment; 6 deletions based on UK comments, and reshuffling of code positions. (The final result is captured in the relevant resolution M48.14 below.)



Disposition: Accept a modified set of net 108 characters, for two technical ballots. Final charts are in document N3059. Name change for LATIN SMALL SCRIPT D to LATIN SMALL LETTER DELTA, and it is moved to 1E9F. See relevant resolution M48.14 below.

*Relevant resolution:*

*M48.14 (108 Medievalist characters):*

*Unanimous*

With reference to document N3027, WG 2 accepts to encode the following in the standard:

- a. Twenty six (**26**) characters in the Combining Diacritical Marks Supplement block from pages 52 and 53 (excluding 1DE4)
- b. Nine (**9**) characters in the Latin Extended Additional block from pages 54 and 55; with the name change for 1E9F to LATIN SMALL LETTER DELTA
- c. Seventy three (**73**) characters in the Latin Extended-D block from pages 58 and 59 (excluding A761 to A766)

The final glyphs, names and code position assignments are reflected in the charts in document N3059.

### **8.11 Mayanist Latin letters**

Input documents:

- [3028](#) Proposal to add Mayanist Latin letters to the UCS; Michael Everson; 2006-01-30  
[3074](#) Expert Input on 'Proposal to add Mayanist Latin letters to the UCS' N3028 (=L2/06-028); Debbie Anderson; 2006-04-07  
[3082](#) Revised proposal to add Mayanist Latin letters; Michael Everson; 2006-04-10  
[3094](#) Discussion of the Mayanist Latin letter CUATRILLO WITH COMMA; Everson; 2006-04-16

Mr. Michael Everson: Document N3028 was the original proposal and is replaced completely with document N3082. Document N3074 is feedback from Mayanist community on the earlier document. The new document addresses all the comments from the community. The glyphs are all improved. Two new characters are requested - document N3094 gives additional justification for these two. Eight of these characters are also in Irish ballot comments in addition to deleting two.

Discussion:

- a. Mr. Mike Ksar: Is there any objection to removing two and adding eight as requested by Ireland?
- b. Dr. Ken Whistler: US national body is of two minds. It is a little bit more complicated. The official position is not to change content of Amendment 3. However, based on ad hoc-ing since the last US national body meeting, and based on discussions on the email lists, there seems to be a consensus building towards adding the eight characters per Irish comments. As long as US national body has sufficient review opportunity the US should be able to address any issues in the next go around.
- c. Mr. Mike Ksar: The next Amendment 3 ballot will be a four-month FPDAM ballot.
- d. Dr. Ken Whistler: My understanding of the discussion so far is that a consensus is building towards accepting these by US national body. I would not like to stand in its way even though US national body has not officially approved such a change.
- e. Dr. Asmus Freytag: It is normal for national body delegation to be presented with new evidence at WG 2 meetings. Anything we approve here goes on a ballot and therefore allows for national body review.
- f. Mr. Mike Ksar: We accept adding eight characters in N3082. Two characters of the eight are replacements for the two that are requested to be removed by Ireland.
- g. Dr. Asmus Freytag: There is some question on locations of all these eight -- we will ad hoc on these and settle on their final position.
- h. Mr. Michael Everson: There is a proposal to add two more characters. Feedback from Mayanists is reflected in N3082. Text in this document sort of misleading -- these ARE new characters -- does not make clear about a Chinese proposal based on their linguist study from Guatemela. Latin Capital and Small letters HENG are proposed additions. The request is also to have all the ten characters contiguously for convenience of use by Mayanists.
- i. Mr. Mike Ksar: Any objections on adding these?
- j. Dr. Ken Whistler: No objections.
- k. Mr. Mike Ksar: Ad hoc can meet and decide on location of TEN characters. Mr. Michael Everson and Dr. Asmus Freytag met and decided on final positions.

Disposition: Accept for Amendment 3. In Latin extended -D; contiguous range of A726 - A72F is the code position range for the ten Mayanist characters.

New characters from N3082:

A726 LATIN CAPITAL LETTER HENG  
A727 LATIN SMALL LETTER HENG

And, replace (in PDAM3)  
2C6F LATIN LETTER TRESILLO  
2C70 LATIN LETTER CUATRILLO  
with  
A728 LATIN CAPITAL LETTER TZ  
A729 LATIN SMALL LETTER TZ  
A72A LATIN CAPITAL LETTER TRESILLO  
A72B LATIN SMALL LETTER TRESILLO  
A72C LATIN CAPITAL LETTER CUATRILLO  
A72D ATIN SMALL LETTER CUATRILLO  
A72E ATIN CAPITAL LETTER CUATRILLO WITH COMMA  
A72F ATIN SMALL LETTER CUATRILLO WITH COMMA  
with glyphs from N3082,

See relevant resolution M48.2 below.

*Relevant resolution:*

*M48.2 (Changes to Mayanist characters):*

*Unanimous*

With reference to document N3082, and in response to PDAM3 ballot comments in document N3071 from Ireland, UK and USA, regarding Mayanist characters, WG 2 accepts the following:

- a. Add two new characters to the Latin Extended-D block:  
A726 LATIN CAPITAL LETTER HENG  
A727 LATIN SMALL LETTER HENG  
with their glyphs that are similar to those shown in document N3082.
- b. Replace the following two (2) caseless characters in Latin Extended-C block:  
2C6F LATIN LETTER TRESILLO  
2C70 LATIN LETTER CUATRILLO  
with the following eight (8) case-paired characters in Latin Extended-D block:  
A728 LATIN CAPITAL LETTER TZ  
A729 LATIN SMALL LETTER TZ  
A72A LATIN CAPITAL LETTER TRESILLO  
A72B LATIN SMALL LETTER TRESILLO  
A72C LATIN CAPITAL LETTER CUATRILLO  
A72D LATIN SMALL LETTER CUATRILLO  
A72E LATIN CAPITAL LETTER CUATRILLO WITH COMMA  
A72F LATIN SMALL LETTER CUATRILLO WITH COMMA  
with their glyphs that are similar to those shown in document N3082.

### **8.12 Ordbok characters**

Input documents:

[3031](#), [3031-1](#), [3031-2](#)

Proposal to encode characters for Ordbok över Finlands svenska folkmål in the UCS;  
Therese Leinonen, Klaas Ruppel, Erkki I. Kolehmainen, Caroline Sandström; 2006-01-26

Mr. Erkki Kolehmainen: Three phonetic characters for editing Swedish dialects in Finland are found to be lacking. The full set is not being proposed by Finland -- Sweden may propose the full set.

Discussion:

- a. Mr. Mike Ksar: Sweden may make contribution for full set in the future.
- b. Mr. Erkki Kolehmainen: These are not part of UPA. Other characters we use are from IPA. Sweden has input via the fonts being sent in. The contributors are from Finland Research institute.
- c. Mr. Mike Ksar: Do you know who would be representing Sweden in the future? (No).
- d. Dr. Asmus Freytag: Have suggested code positions .. 2C78, 79 and 7A. Name change for 2C7A -- from 'O with Ring Inside Down' to 'O with Low Ring Inside'.

Disposition: Accept three characters from N3031 in Latin extended C; with one name change. See relevant resolution M48.21 below.



*Relevant resolution:*

*M48.21 (3 Phonetic characters for Swedish dialects in Finland):*

*Unanimous*

With reference to document N3031, WG 2 accepts the following three (3) characters to the Latin Extended-C block:

2C78 LATIN SMALL LETTER E WITH TAIL  
2C79 LATIN SMALL LETTER TURNED R WITH TAIL  
2C7A LATIN SMALL LETTER O WITH LOW RING INSIDE (changed name)  
with their glyphs that are similar to that shown in document N3031.

### **8.13 Addition of two Mathematical brackets**

Input document:

[3040](#) Proposal for the Addition of Two Mathematical Bracket Characters; Unicode/USA NB (contact: Asmus Freytag, Barbara Beeton, Murray Sargent); 2006-02-09

Dr. Asmus Freytag: A couple of years ago, we took a WG 2 decision to separate out brackets used for CJK from those in Math due to differences in their font properties. At that time we did not have attestations for White Tortoise Shell Brackets. Since then we have double checked the collections of Math characters, and as new fonts are developed for Math, these have become obvious candidates for being added to an existing set, continuing the disunification with the CJK brackets.

Dr. Ken Whistler: The US national body has reviewed this and are supporting this.

Disposition: Accept the two characters from document [N3040](#). See relevant resolution M48.22 below.

*Relevant resolution:*

*M48.22 (2 Mathematical Brackets):*

*Unanimous*

With reference to document N3040, WG 2 accepts the following two (2) characters in the Miscellaneous Mathematical Symbols-A block:

27EC MATHEMATICAL LEFT WHITE TORTOISE SHELL BRACKET  
27ED MATHEMATICAL RIGHT WHITE TORTOISE SHELL BRACKET  
with their glyphs that are similar to those shown in document N3040.

### **8.14 Arabic Math characters**

Input documents:

[3085](#), [3085-1](#) Arabic Mathematical Alphabet Symbols; .doc: Summary Proposal Form; .pdf: Supporting material; Azzeddine LAZREK; 2006-03-30

[3086](#), [3086-1](#) Diverse Arabic Mathematical Symbols; .doc: Summary Proposal Form; .pdf: Supporting material; Azzeddine LAZREK; 2006-03-30

[3087](#), [3087-1](#) Rumi Numeral System Symbols; .doc: Summary Proposal Form; .pdf: Supporting material; Azzeddine LAZREK; 2006-03-30

[3088](#) The others Arabic Mathematical Symbols that are not accepted; Azzeddine LAZREK; 2006-03-26

[3089](#) A complete Rumi bibliography; Azzeddine LAZREK; 2006-03-31

(Note: In the above set, documents numbered with the -1 have the character proposals; the corresponding document without the -1 in its number has only the proposal summary form.)

Dr. Asmus Freytag: We knew that typesetting of Math in Arabic needed special considerations. We did not have enough evidence and attestations about these in the past. Now we have some input. The UTC has worked with the author on refining the proposals. I will present these proposals on behalf of the author Mr. Azzeddine Lazrek.

Referring to document [3086-1](#) -- Symbols for the square root - we have in the standard symbols for cube and fourth root etc. In Arabic, instead of using the European Numerals with root sign --- they use the right to left version of the symbols with Arabic-Indic Numerals instead of Western digits. The proposal is for Arabic-Indic Cube Root and one for Arabic-Indic Fourth Root.

Discussion:

- a. Mr. Michael Everson: Will there be similar requirement for similar characters with Persian-Indic digits? For example Roozbeh could be contacted.
- b. Dr. Asmus Freytag: If they can demonstrate the requirement we could consider those.

Dr. Asmus Freytag: There is also a Letter-Like Symbol called Arabic Ray. We have per cent and per mille in western ... in Arabic per mille and per ten thousand, the zeros are replaced with Arabic-Indic zeroes (dots). Two characters are proposed for this. Outlined White Star, which is a five-pointed star, and which is different from a pentagram. Section 6 proposes a set of Mathematical Arrows. In Arabic Math use when using Right to Left text layout, all the Arrows should point in a direction opposite to the

use of the arrows in Western Math. In Math one could write relations in reverse direction conceptually. Normally these should not matter because their mirrored property is used to get the desired results. However, we have deliberately excluded the Arrows from mirrored property. This leads to these being requested additionally.

Dr. Umamaheswaran: We should state explicitly in the standard about a statement that Arrows will not be added to the list of characters having a mirroring property.

Dr. Asmus Freytag: Referring to document [N3085-1](#) -- particular shapes of letters are used in Math to associate with particular meanings. In Western Math we added a large collection - what looks like styled Alphabets in Latin set. The Table 2 in this document on page 5 contains a list of such characters in Arabic Math. Some of the characters in this Table 2 are in the presentation forms of Arabic in the standard. We are in dialog with the author on that score. Any missing characters are candidates for inclusion in the standard. Table 3 - shows particular symbols used in Arabic Math -- like Letter-Like symbols of Western Math. These cannot be left to Fonts / shaping systems to handle. Table 4 - is a collection of Double-struck forms. Section 4 - lists a number of Dotless Letter-like symbols. Section 5 -- frozen Math forms for alphabetic symbols. Section 6 - has a number of operators etc. In Western Math we use Greek Letters --- for example, large sigma for summation. Arabic Math have similar large symbols -- in dotted and dotless forms. Western summation symbol grows in height, whereas the Arabic summation symbol grows in Length. An example is shown on page 11. It is typographically different from the Arabic word for Sum. The rest of the document has a number of examples.

Document [N3087-1](#) --- Rumi Numeral System Symbols. - is the Traditional / historical Arabic numbering system used in North Africa -- the RUMI system. The Unicode consortium had requested for more information. Document [N3089](#) is a complete bibliography of the available documents. We have not studied these documents to make any recommendation at this time. We need to get someone who can read these references and make some judgement. The proposed set seems to be mature... but our understanding of use of these characters in the system needs more study.

Dr. Ken Whistler: The current request is for 31 characters that include nine singles, nine tens, nine hundreds and four fractions symbols.

Discussion:

- a. Mr. Michael Everson: How many of these characters are requested?. Can we fill the current empty space in presentation symbols? I have no objection to putting them in a separate Arabic Math block.
- b. Dr. Ken Whistler: There are 200 total characters proposed in total -- between the three proposal documents. These proposals are not quite ready for ballot. The evidence for these characters is compelling -- and most of these should go into the standard. The next step is to rationalize these and separate them into the presentation block, special Math block etc.
- c. Dr. Asmus Freytag: There is a subset of these that could be included in the presentation forms area. We don't have a position on the code positions. In terms of making progress either WG 2 or UTC has to take ownership of the tables. The contributors have given us as much information as they can.
- d. Mr. Mike Ksar: We acknowledge the contribution; it is worthwhile pursuing these proposals.

Action item: US national body is invited to work with the authors and come up with a proposal that could be included in the standard. Other national bodies to review and feedback.

### **8.15 Mongolian – Manchu Ali Gali letter**

Input document:

[3041](#) Proposal to encode one Manchu Ali Gali letter - Mongolian; Andrew West; 2006-02-13

Mr. Mike Ksar: The proposal from Mr. Andrew West. I have a note from China that they support this proposal. The UK is back as P member of SC 2.

Dr. Ken Whistler: US national body also supports this proposal.

Disposition: Accept. See relevant resolution M48.23 below.

*Relevant resolution:*

*M48.23 (1 Mongolian character):*

*Unanimous*

With reference to document N3041, WG 2 accepts the following character in the Mongolian block:

18AA MONGOLIAN LETTER MANCHU ALI GALI LHA  
with its glyph similar to that shown in document N3041.

## **8.16 Myanmar additions**

### **8.16.1 Additional 7 Myanmar characters disunified**

Input documents:

- [3043](#) Proposal to encode seven additional Myanmar characters in the UCS; Ireland (NSAI), United Kingdom (BSI), Myanmar Language Commission, Myanmar Unicode and Natural Language Processing Research Center, Myanmar Computer Federation; 2006-03-07
- [3061](#) Further feedback on n3043 - Myanmar; Martin Hoskin; 2006-03-27
- [3065](#) Additions to AMD 3 - Tibetan, Mongolian, Myanmar, and Lao; UK NB - Andrew West; 2006-03-27
- [3069](#) Concerns Regarding WG2 N3043, Myanmar Additions to 10646; Unicode Consortium; 2006-04-06
- [3078](#) Proposed additions to 10646 Amendment 3; Irish NB; 2006-04-06
- [3079](#) Response to UTC contribution N3069 "Concerns on N3043R, Nyanmar Additions"; Irish and UK NBs; 2006-04-08
- [3083](#) A Sgaw Karen Unicode Proposal; Extending Myanmar to Incorporate Sgaw Karen; Martin Hosken, SIL International and Payap University; 2006-04-10
- [3084](#) Dissociating Myanmar Medials; A Proposal to Encode Separate Myanmar Medials; Martin Hosken SIL International and Payap University; 2006-04-10

Output document:

- [3099](#) Myanmar ad hoc report - meeting 48; Ad hoc - Erkki; 2006-04-26

Mr. Mike Ksar: We should have an ad hoc to consider and discuss these contributions. I expect a recommendation from the ad hoc group for discussion at the main meeting.

The ad hoc group participants were Mr. Erkki Kolehmainen (chair), Mr. Michael Everson, Dr. Asmus Freytag, Ms. Deborah Anderson, Ms. Lisa Moore, Mr. Thein Oo, Mr. Tseng Shih-Shyeng, Ms. Wei Lin-Mei, Dr. Ken Whistler and Mr. Zaw Htut.

Mr. Erkki Kolehmainen: The ad hoc met and its conclusions are in the ad hoc report in document [N3099](#). The recommendation is to include the seven characters proposed in document [N3043](#). Current specifications in the standard does not permit a number of things -- we cannot represent all the languages written using Myanmar script. We also cannot represent distinctions between words that have to be made. It requires a change in the encoding model. A positive side effect is that Burmese language encoding also is made simpler even though that is NOT the main reason for the proposal. The ad hoc was in consensus with the recommendations. We should encode seven characters in [N3043R](#) -- with the resultant implications. It should be for a two technical ballots. Ad hoc focused on all the Myanmar documents except [N3044](#). It also addressess requests for additions to Myanmar contained in other national body comments accompanying the Amendment 3 ballot.

Discussion:

- a. Mr. Mike Ksar: I thank Mr. Erkki Kolehmainen and ad hoc group in discussing and arriving at this consensus.
- b. Dr. Umamaheswaran: There are two glyph corrections in document N3073.
- c. Dr. Ken Whistler: The second glyph correction in document N3043 is non controversial. The first one -- for Virama -- is the consequence of the acceptance of the seven new characters.

Disposition: Per ad hoc report, accept proposed seven characters for encoding from document [N3043R](#), disunifying from current encodings for these. See relevant resolution M48.15 below.

*Relevant resolution:*

*M48.15 (7 Myanmar additions):*

*Unanimous*

With reference to documents N3043 and N3099, WG 2 accepts the following changes in the Myanmar block:

Add the following seven (7) characters:

102B MYANMAR VOWEL SIGN TALL AA  
103A MYANMAR SIGN ASAT  
103B MYANMAR CONSONANT SIGN MEDIAL YA  
103C MYANMAR CONSONANT SIGN MEDIAL RA  
103D MYANMAR CONSONANT SIGN MEDIAL WA  
103E MYANMAR CONSONANT SIGN MEDIAL HA  
103F MYANMAR LETTER GREAT SA

with their glyphs that are similar to those shown in document N3043.

Change the glyphs for the following:  
104E MYANMAR SYMBOL AFOREMENTIONED  
1039 MYANMAR SIGN VIRAMA  
to the glyphs that are similar to those shown in document N3043.

## 8.16.2 Myanmar additions for Mon and S'gaw Karen languages

Input document:

[3044](#) Proposal for encoding Mon and S'gaw Karen characters in the UCS; Michael Everson and Martin Hosken; 2006-04-08

Mr. Michael Everson: When self and Martin Hosken were in Yangon (formerly Rangoon), the model change has to be accepted first before minority languages can be accommodated. Since we have accepted the model change, 14 characters in support of two minority languages Mon and S'gaw Karen are proposed. The latter language was discussed in detail in the ad hoc in model discussion. One dependant vowel, two independent vowels, and some consonants required for Mon and one independent letter and three signs for S'gaw Karen are proposed. There were empty slots in the Myanmar block. We propose to fill some of these with Mon and S'Gaw Karen characters. There are also other proposals.

Discussion:

- a. Dr. Ken Whistler: It was known at the time we had made the Myanmar block that there were several other languages that were yet to be researched in the London meeting. The first of the S'gaw Karen looks like a combination of two characters RA + xxx .. at 1061 - SHA. Users prefer that it be realized as a single character - it has a sorting position totally different from ra. Figure 9 shows its relative collation position. Figure 10 is another view of the alphabetical order. The US national body has not had a chance to look at these. My personal opinion is that this collection is well documented and justified. We can put these into an Amendment when we have two rounds of technical ballots.
- b. Mr. Yasuhiro Anan: There are other Karens.. on the addition of SHA .. can be represented by other characters. The other Karens have different similar characters.
- c. Mr. Michael Everson: the intent is to use the same SHA character for all Karens. 1063 and 1064 should have the words 'tone marks' in their names.
- d. Mr. Mike Ksar: Dr. Ken Whistler had chaired the Myanmar ad hoc in London. Mr. Yasuhiro Anan had chaired the Toronto and Xiamen ad hocs in Myanmar. I also thank the experts from Myanmar for their contributions.

Disposition: Accept 14 Myanmar characters from document N[3044](#); two characters with name changes. See relevant resolution M48.16 below.

*Relevant resolution:*

*M48.16 (14 Myanmar characters for Mon and S'gaw Karen):*

*Unanimous*

WG 2 accepts to encode in the Myanmar block the fourteen (14) characters, some of which are combining, with their glyphs that are similar to those shown in document N3044:

1028 MYANMAR LETTER MON E  
1033 MYANMAR VOWEL SIGN MON II  
1034 MYANMAR VOWEL SIGN MON O  
105A MYANMAR LETTER MON NGA  
105B MYANMAR LETTER MON JHA  
105C MYANMAR LETTER MON BBA  
105D MYANMAR LETTER MON BBE  
105E MYANMAR CONSONANT SIGN MON MEDIAL NA  
105F MYANMAR CONSONANT SIGN MON MEDIAL MA  
1060 MYANMAR CONSONANT SIGN MON MEDIAL LA  
1061 MYANMAR LETTER SGAW KAREN SHA  
1062 MYANMAR LETTER SGAW KAREN EU  
1063 MYANMAR TONE MARK SGAW KAREN HATHI  
1064 MYANMAR TONE MARK SGAW KAREN KE PHO

## 8.17 Saurashtra script

### 8.17.1 Correction of a Saurashtra character name in PDAM3

Input documents:

[3058](#) Correction of a Saurashtra character name in PDAM3; Everson; 2006-03-23

Mr. Michael Everson: Dr. Peri Bhaskara Rao had criticized the name Upakshara for one of the characters. Based on further research on work by Mr. H.N. Randle, there are three marks of the

category Upakshara -- the name of A8B4 is proposed to reflect this to be HAARU instead of Upakshara. In addition to changing to Haaru, Ireland in its PDAM3 ballot comments further requests that the name for A8B4 be changed to - Saurashtra Consonant Sign Haaru.

Discussion:

- a. Mr. Mike Ksar: Any objections?
- b. Dr. Ken Whistler: I have no objection.

Disposition: Accept the name change modified per Irish ballot comments. It is still a combining character. See relevant resolution M48.4 below.

*Relevant resolution:*

*M48.4 (Change name of one Saurashtra character):*

*Unanimous*

With reference to document N3058 and Irish PDAM3 ballot comment T.4 in document N3071, WG 2 accepts to change the name of A8B4 in the Saurashtra block:

from SAURASHTRA LETTER UPAKSHARA  
to SAURASHTRA CONSONANT SIGN HAARU.

### 8.17.2 Implementation of the Saurashtra script in N2969

Input documents:

[3076](#) Implementation of the Saurashtra script (N2969R2) (=L2/05-222R2) ; Debbie Anderson; 2006-04-10

This document describes how the current Saurashtra encoding has been implemented. It is for information of the delegate. There was no discussion.

## 8.18 Lithuanian

### 8.18.1 Named sequences and collection for Lithuanian

Input document:

[3090](#) Proposal to identify the Lithuanian Alphabet as a Collection in the ISO/IEC 10646, including the named sequences for the accented letters that have no pre-composed form of encoding (also in TUS; Erkki I. Kolehmainen (member of CEN/ISSS CDFG), based on input from the Lithuanian Language Institute; 2006-04-12

Output document:

[3105](#) Extended Collection definition; Michel Suignard; 2006-04-27

Mr. Erkki Kolehmainen: A Lithuanian request for encoding composed characters separately has been there for a long time. However, the normalization will not let this happen. Lithuania came to (Focus Group on Cultural Diversity) CDFG requesting CDFG to take a resolution to demand the separate encoding from WG 2. I dissuaded them - and since CDFG was not a technical committee we cannot dictate a technical solution. This contribution is to help Lithuanian to get them what they would need. The proposal is to have named sequences for the combined characters that could be identified in the Lithuanian collection and their standards. Addendum I of document shows the accented letters that are already in the standard -- some as fully composed. Others can only be sequences. Addendum II has identified those that are sequences with names etc. Addendum III shows the Lithuanian standard repertoire. The editor has enough information in Addendum I.

Discussion:

- a. Mr. Mike Ksar: The request is to have a list of named sequences and one or more collections. Addendum I defines the collection.
- b. Mr. Michel Suignard: Collection can contain sequences?
- c. Dr. Asmus Freytag: Currently USIs could be included in collections.
- d. Mr. Yasuhiro Anan: Japan has a contribution and have to check on whether collection can contain USIs. Clause 6.6 states USIs may not be used in collections etc.
- e. Dr. Ken Whistler: On the addition of USIs -- take input from Addendum II (Minus those that are already in the named sequences list).
- f. Mr. Mike Ksar: Need to check if a collection can contain a USI as one of its elements.
- g. Mr. Michel Suignard: USIs were meant to be used outside the context of ISO/IEC 10646.
- h. Dr. Ken Whistler: If there is consensus that a collection can contain a USI then we may need to take a separate resolution to remove the note from the current text in the standard.
- i. Dr. Ken Whistler: The intent was to allow USIs and let others use these in their standards / definitions. In this case Lithuanian standard can include the USIs. The problem here is that Lithuanian national body is unable to see it as a sufficient solution for their problem. It is not sufficient for them to have it in their document - it should be in ISO/IEC 10646. If we entertain this request, ISO/IEC 10646 would become potentially a register of



- collections required by other standards. As long as it does not set a precedent we may be able to entertain and accommodate the Lithuanian request.
- j. Mr. Mike Ksar: I recall the discussion now. We did not want to become a registration authority for others.
  - k. Mr. Michel Suignard: Japan had used the methodology -- in their national standard they have defined the needed sequences in their standard.
  - l. Dr. Ken Whistler: Japan should be the precedence. However, Lithuanian national body keeps coming back with a request for separately coded characters. If we can accommodate their request it will satisfy them.
  - m. Mr. Mike Ksar: Can we approach Lithuania and communicate this to them.
  - n. Mr. Erkki Kolehmainen: Lithuania will certainly come back through European Union.
  - o. Mr. Michel Suignard: The current collection definition is normative, the collections are also normative.
  - p. Dr. Ken Whistler: Alternative is to have extended collections -- without disturbing the current collections definitions. This is an international standard with political ramifications as well. If an annex can be used to satisfy national bodies then we should strive.
  - q. Dr. Umamaheswaran: I am proposing the outline of three possible resolutions.
    - a. Add Named sequences from Addendum II after eliminating any duplicate from existing named sequences list in clause 29 Amendment 2. (Only one duplicate no. 14 in Addendum II was identified by Mr. Erkki Kolehmainen.)
    - b. Define an 'extended collection' -- that permits use of USIs in them.
    - c. Information from Addendum I can then be defined as an 'extended collection'.
  - r. Mr. Yasuhiro Anan: I would support the direction -- would like to check with others.
  - s. Mr. Mike Ksar: We have consensus on the direction, and we will approve the text of the resolution.

The project editor had prepared document [N3105](#) containing a proposed definition for Extended Collection. This was briefly discussed prior to the relevant resolution being adopted.

Prior to adoption of the relevant resolution a problem was identified. Dr. Umamaheswaran outlined the problem on the white board. A combined character like A with Ogonek with Acute - entry 2 in Addendum I - one of the requested named sequences in document N3090, has a canonically equivalent named sequence A with Acute and Ogonek in the standard - clause 29 Amendment 2. There are four possible sequences that are all canonically equivalent ... A-acute with Ogonek, A with Acute with Ogonek, A-ogonek with Acute, A with Ogonek with Acute. Of these the last one is the fully decomposed normalization form NFD and the third one is the NFC form. The current standardized named sequence is neither of these normalization forms. There is a lack of clear rules / guidelines on the naming of Named Sequences, and what to do with canonically equivalent sequences. The standard has currently a 'name uniqueness rule'.

It was decided that the project editor is given instructions to come up with some rules and these can be reviewed and revised during the ballot review and resolved at the next meeting.

Dr. Asmus Freytag: We don't want to end up in ignoring customer-preferred names like the situation with character names, leading to 'misspelled character names list'. Even for the same sequence different user communities may have different names.

Action item: Input into principles and procedures -- extend definition to include extended collections and format for submitting such a request. The requested sequence A with Ogonek with Acute will be included in the standard. The editor will adjust the names for these Named Sequences in line with the rules he will be including in the Amendment.

Disposition: The Lithuanian request for named sequences is accepted; a new 'extended collection' definition will be added to accommodate the Lithuanian collection. See relevant resolutions M48.26, M48.27 and M38.28 below.

*Relevant resolutions:*

*M48.26 (Named sequences for Lithuanian):*

*Unanimous*

With reference to document N3090, WG 2 accepts the named sequences for composite characters for Lithuanian repertoire identified in the table in Addendum II less the following:

Entry 14, which is already encoded as LATIN SMALL LETTER I WITH DOT ABOVE AND ACUTE, 0069 0307 0301 (identical to the requested sequence) in clause 29 Amendment 2 to the standard. WG 2 further instructs its project editor to add text to the standard to clarify the rules for adding and naming Named Sequences, with particular attention to those resulting in canonically equivalent composed characters.

*M48.27 (Extended Collections):*

*Abstention: Japan*

*Approval: Canada, Finland, Ireland, Korea (Republic of), Poland and the USA*

WG 2 accepts the definition of 'Extended Collection' given in document N3105 for inclusion in the standard.

*M48.28 (Lithuanian Extended Collection):*

*Unanimous*

WG 2 accepts the inclusion in the standard of an extended collection named 'Contemporary Lithuanian Letters', an extended collection of standalone characters and named sequences, corresponding to the repertoire shown in Addendum I in document N3090.

## 8.18.2 Two combining characters for Lithuanian phonetics

Input document:

[3048](#)

Proposal to encode two combining characters in the UCS; State Commission of the Lithuanian Language (contact: Vldas Tumasonis); 2006-03-02

Mr. Erkki Kolehmainen: Document [N3048](#) is a request for two combining characters. Historical origin is to have combination of Breve and Macron and Macron and Breve.

Discussion:

- a. Dr. Ken Whistler: We have done similar combining marks recently -- circumflex with grave, with acute etc. The proposed characters are in line with that; the rationale provided is sound.
- b. Mr. Mike Ksar: Are these part of the named sequences in the other document? No.
- c. Dr. Ken Whistler: These characters are used by Linguists for representing Lithuanian sounds / phonetics; they are not part of Lithuanian alphabet.

Disposition: Accept with names modified. See relevant resolution M48.17 below.

*Relevant resolution:*

*M48.17 (2 combining marks for Lithuanian phonetics):*

*Unanimous*

With reference to document N3048, WG 2 accepts the following two (2) combining marks in the Combining Diacritics Supplement block:

1DCB COMBINING BREVE-MACRON

1DCC COMBINING MACRON-BREVE

with their glyphs that are similar to those shown in document N3048; both are combining characters.

## 8.19 CJK Strokes additions

Input document:

[3063](#)

Proposed additions to the CJK Strokes block of the UCS 3. IRG rapporteur's summary report; Lu Qin - IRG; 2006-04-03

Dr. Lu Qin: The IRG was instructed to consolidate the various proposal for CJK strokes. The result is that we are proposing 20 additional stokes in the CJK Strokes block. A table of code positions and names is given, along with where the stokes are used. HKSCS had some strokes in it. To make HKSCS complete the stroke block was completed -- there was a proposal to include the complete stroke collection. A proposal summary form is also attached. Except for the last one, all the others are recognizably Chinese -- the last one is submitted by Korea.

Discussion:

- a. Mr. Yasuhiro Anan: There is an example of use of all these strokes -- the last one shows the use of the circular stroke. The first example is a CJK Unified ideograph,
- b. Mr. Michel Suignard: When we were looking at the CJK Extension B, there was a cursive stroke character that looks like a Greek Alpha - a stroke inside 201AD.
- c. Mr. Yasuhiro Anan: The proposal includes only the strokes on which there are no questions.
- d. Dr. Lu Qin: If there are new characters discovered using unseen strokes before, we may have to extend the set. We know that 201AD is not covered.
- e. Dr. Asmus Freytag: We got a report that the said character seems to be in the wrong font because it is cursive.
- f. Dr. Ken Whistler: On the collation stuff, these are presented in the LDML format. What is not clear is whether these should be included in the CTT of ISO/IEC 14651 or only for tailoring using CLDR.



- g. Mr. Yasuhiro Anan: We have provided the collation data because of the submission guidelines. We followed the CLDR compatible format. It is not our intention to submit to CLDR.
- h. Dr. Ken Whistler: I am trying to get to the next step -- if anything is going to happen with the collation data, either WG 2 or submitters have to do something with it. Either ignore it -- which means it ends up in a binary order. Or the ISO/IEC 14651 editing group could be requested to use the order presented here. Or let the binary order stand and request whoever cares about the order out of IRG can ask CLDR group to accommodate it in tailoring.
- i. Mr. Mike Ksar: Thanks for pointing this out. I am inclined to say we should forward this to attention of ISO/IEC 14651 group. The reason we have collected this information is to help with ISO/IEC 14651.
- j. Dr. Ken Whistler: It is appropriate for WG 2 to make some recommendation about this section.
- k. Mr. Erkki Kolehmainen: We cannot submit to CLDR we have to await acceptance of this character.
- l. Dr. Lu Qin: The naming convention that was used is outlined in the document -- following the principles used before.

Disposition: Accept the 20 CJK strokes as in document [N3063](#). See relevant resolution M48.12 below.

Action item: Request SC 2 to forward document [N3063](#) to ISO/IEC 14651 editing group to accommodate the tailoring of order in a future Amendment to ISO/IEC 14651.

*Relevant resolution:*

*M48.12 (CJK Strokes):*

*Unanimous*

With reference to document N3063, WG 2 accepts to encode twenty (20) additional characters in code positions 31D0 to 31E3 in the CJK Strokes block, with names, glyphs and code position assignments as shown on page 2 of document N3063.

## **8.20 Phaistos Disc script**

Input document:

[3066](#) Phaistos Update; Everson, Jenkins; 2006-04-01

Mr. Michael Everson: Document [N3066](#) replaces the earlier preliminary proposal document [N1575](#) of 1997. Since the initial proposal there has been great deal of interest on Phaistos disc script and great deal of interest on what is written on it. A list of references is given in document. The names describe the characters. If it is a writing system it has not been established. It is being treated as a writing system -- currently a couple of private use area implementations. In PUA we cannot do global searching etc. possible. Because of increased interest in this and to assist the researchers we would like to encode these in the SMP. Self and John Jenkins are proposing to include these in the standard to assist the interested people.

Discussion:

- a. Dr. Ken Whistler: there are decipherers. There are no native users for almost 3000 years. There are a number of users interested in deciphering these; they are using these as symbols. US national body had reviewed the old version of the document and the assessment at that time was that this was an undeciphered script and is not a candidate. We have not reviewed this new proposal. My initial assessment is that the first 45 of these are all over the place and can be used as enumerated symbols. The 46th one is also seen on the disk, but not as a symbol. The last two are artefacts of layout of these things -- they are not signs, nor punctuations etc. However, the set as a whole should be used as a set of symbols to help the decipherers.
- b. Mr. Yasuhiro Anan: When these symbols do change the nature from symbol and get deciphered to a letter does it affect the character properties?
- c. Dr. Ken Whistler: It is a question for Unicode not for WG 2.
- d. Dr. Asmus Freytag: The distinction is not that much of a problem.
- e. Dr. Ken Whistler: My personal opinion is that this proposal -- minus the last two characters -- is ready to go ahead. The other 45 plus the sign are not likely to get any better even if we wait for next ten years or so.
- f. Dr. Asmus Freytag: We can take it as is -- request feedback. Or we can put it on the ballot with some modifications, and wait for two technical ballots.

- g. Mr. Michael Everson: When I used these in the PUA users have asked for the last three characters. I agree that the last two vertical bars could be some others from the standard. The combining mark is there for some use. I agree we can remove the last two.
- h. Dr. Asmus Freytag: If we want to create some ancient script general punctuation mark I may be agreeable.
- i. Dr. Ken Whistler: When we draw something on a circle (the disc) you show vertical lines to separate segments -- but these are not appropriate as characters. We can use vertical bars in the standard for the same purpose.
- j. Dr. Umamaheswaran: I think we should go ahead with the proposal with 46 characters (one combining mark).

Disposition: Accept per document N[3066](#) excluding 101FE and 101FF. See relevant resolution M48.13 below.

*Relevant resolution:*

M48.13 (Phaistos disc symbols):

*Unanimous*

WG 2 resolves to encode the Phaistos disc symbols proposed in document N3066 in a new block 101D0 - 101FF in the SMP, named 'Phaistos Disc', and populate it with forty six (46) characters (one of which is combining) with the names, glyphs and code position assignments as shown on pages 14 and 15 (excluding two characters at 101FE and 101FF) in document N3066.

**8.21 Three Uralic phonetic additions**

Input document:

[3070](#) Proposal for Encoding 3 Additional Characters of the Uralic Phonetic Alphabet; Klaas Ruppel, Jack Rueter, Erkki

Mr. Erkki Kolehmainen: The Uralic Phonetic Alphabet is continuously evolving. In the initial submissions we included those for which we had solid proof of use. We are continuously digging up the stuff we need. We added a few in Sophia Antipolis meeting that we needed right away. We have found additional characters in actual use in the material being researched and being encoded on computers for research and publications purposes. There is an addendum that explains the nature of the UPA and the need to come with more requests for encoding. The current request is for three attested ones. More may come in the future.

Discussion:

Mr. Mike Ksar: I note that UPA is an evolving one.

Dr. Asmus Freytag: Table on page 3 in N[3070](#) shows the proposed three characters. Code points have to be changed from those proposed. The new code positions are 2C7B--2C7D. I will need the font for these.

Mr. Erkki Kolehmainen: The note from Klaas Ruppel to Mr. Mike Ksar in April had the font, font license was also included. Mr. Erkki Kolehmainen will forward to Dr. Asmus Freytag the fonts and licence.

Disposition: Accept three UPA characters as in document N3070; new code positions. See relevant resolution M48.19 below.

*Relevant resolution:*

M48.19 (3 Uralic Phonetic characters):

*Unanimous*

With reference to document N3070, WG 2 accepts the following three (3) characters in the Latin Extended-C block:

2C7B LATIN LETTER SMALL CAPITAL TURNED E

2C7C LATIN SUBSCRIPT SMALL LETTER J

2C7D MODIFIER LETTER CAPITAL V

with their glyphs that are similar to those shown in document N3070.

**8.22 Support for OI Chiki script**

Input document:

[3075](#) Support for OI Chiki (N2984R =L2/05-243R); Debbie Anderson; 2006-04-10

For information of the delegates. The document has signed letters from several user groups of Santhali language supporting the encoding of OI Chiki as it is document in N[2984](#) (and has been standardized based on that).

### 8.23 Karen, Shan and Kayah characters

Input document:

[3080](#) Preliminary proposal for encoding Karen, Shan and Kayah characters; Michael Everson and Martin Hosken; 2006-04-09

This document is a preliminary proposal. There was no discussion at this meeting.

Action item: National bodies are to review and feedback.

### 8.24 Vai additions

Input document:

[3081](#) Proposal for addition of Vai characters; Michael Everson, Mohamed Nyei, Charles Riley, Tombekai Sherman; 2006-04-09

Mr. Michael Everson: Document [N3081](#) provides the rationale for request by Ireland in its ballot comment T.3 for adding new Vai characters. It includes Ten digits that are not in the Amendment, but now attested. Nasalized vowels are also included. One of the languages has nasalized vowels - without separate encoding we cannot represent these sounds. These are additional characters in the Irish ballot comments. See reorganized table as part of Irish ballot comments in document [N3071](#).

Discussion:

- a. Mr. Mike Ksar: Any comments?
- b. Dr. Ken Whistler: The US national body has not seen this contribution -- presented now. Personally speaking, I had some discussions and seen the contribution. The case of the digits is clear. These should be added. The case for Nasal Vowels (first four) - the argument made by Mr. Michael Everson for filling in the gaps in the table currently under ballot, with the premise for borrowed vocabulary in languages, is a solid argument for including these.

Disposition: Accept 14 new characters - as proposed in Irish ballot comments. This will reorganize the code chart. It will add a new column to the Vai block. The collection definition has to be adjusted. The content of the block is reorganized. See relevant resolution M48.3 below. See final chart in document [N3059](#).

*Relevant resolution:*

*M48.3 (Changes to Vai script):*

*Unanimous*

With reference to document N3081 and Irish PDAM3 ballot comment T.3 in document N3071, WG 2 accepts to replace the existing Vai block with the charts and names lists as shown in document N3059, which include the following changes:

- a. extending the Vai block by 1 column to end at A62F
- b. addition of the following fourteen (**14**) characters to the Vai block:  
Ten Vai Digits - Zero through Nine and Four new Vai syllables EEN, IN, OON and UN.

### 8.25 Mahjong symbols

Input document:

[2975](#) Proposal to add Mahjong symbols to ISO/IEC 10646; TCA; 2005-08-22

Mr. Tseng Shih-Shyeng: The proposal has been revised based on input from the last meeting (and is posted on the WG2 site). On page 8 there is a blank back. A new section on how to play - we describe two methods to play Mahjong in Taiwan. They are 16 tile game and 13 tile game. We provide more evidence of playing on web sites. We also give examples of books on how to play Mahjong. Other examples are also shown. Answering Dr. Ken Whistler's question, colour does not have any difference but only the pattern matters.

Discussion:

- a. Mr. Mike Ksar: How many characters? 43.
- b. Dr. Asmus Freytag: Your proposal uses jpg files for images -- correct? Yes. The editors cannot progress these without having a True Type Font for these. For simple stuff we could fill in the gap -- but for this set, as an editor I have to argue against progressing this till I get the font. Otherwise, I don't see any particular technical obstacle to encoding these.
- c. Mr. Mike Ksar: Are there suggested code positions?
- d. Mr. Michel Suignard: There is one symbol shown in colour -- we cannot use the colours. The names need adjustment like adding Mahjong Tile Symbol as prefix.

- e. Dr. Ken Whistler: I think for the flowers and seasons it is appropriate for a character encoding standard to have these kind of pictures in the standard. We are encoding 43 symbols for the tile set, not 43 pictures for the tile set. These pictures may be good for someone to sell the tiles ... but they should contain only a symbol not the picture. If we put the picture along with the symbols we cannot read it. The same with the rooster -- we should keep only the symbol and not the picture of Rooster. Once we modify without pictures then there is a chance to create a font. The symbol character will be on each title -- the pictures could change.
- f. Mr. Michel Suignard: At small point size we can have a big picture -- but a large point size will not have room for a picture.
- g. Dr. Asmus Freytag: On page 8, for the Pies -- we have to consider whether a minimum representation should have the details of the picture for the pies or not. In small point sizes it will end up being a dot.
- h. Mr. Mike Ksar: Were these pictures in the previous contribution? How come we did not comment on these?
- i. Dr. Asmus Freytag: It was discussed briefly.
- j. Mr. Michael Everson: It would be a candidate for SMP. Needs more refinement.
- k. Mr. Mike Ksar: You have feedback related to the fonts, representative shapes etc. You don't have to consider the code positions right now. The font policy was established a few meetings ago.

Action item: TCA to take the feedback and produce a new contribution.

## 8.26 Cyrillic combining marks

Input documents:

[3097](#) Proposal to add old Cyrillic titlo-letters to the UCS; UTC + Alexey Kryukov, Vladislav Dorosh; 2006-01-30

Dr. Asmus Freytag: UTC had received a proposal for combining Cyrillic marks in document N3097. These were accepted by UTC. However, there has been some feedback on them that may cause UTC to either change them or even withdraw these.

Action item: National bodies to review document N[3097](#) and feedback. (Document N[3059](#) draft containing the charts for the amendment will be updated to exclude these.)

## 8.27 Proposed additions to Amendment 3

### 8.27.1 Proposals from US national body

Input document:

[2241](#) Proposal to add 6 Egyptological characters to the UCS; Everson  
[3045](#) Additions to ISO/IEC 10646:2003; US National Body; 2006-02-22  
 N3029

Document N[3045](#) contains nine items proposed for Amendment 3. All items except for items 4 and 6 are covered by other agenda items.

- item 1. Clarification concerning Hangul syllable method - see section 9.1 on page 34.
- item 2. New Malayalam characters in document N[3015](#) - see section 8.3 on page 16.
- item 3. New Tibetan characters from documents N[3011](#) and N[3012](#) - see sections 8.1.2 and 8.1.3 on page 15.
- item 5. New Gurmukhi character in document N[3021](#) - see section 8.6.1 on page 17.
- item 7. New Latin characters for Ordbok in document N[3031](#) - see section 8.12 on page 21.
- item 8. New bracket characters in document N[3040](#) - see section 8.13 on page 22.
- item 9. New Mongolian character in document N[3041](#) - see section 8.15 on page 23.

#### 8.27.1.1 Latin Characters for Egyptian transliteration

Input document:

[2241](#) Proposal to add 6 Egyptological characters to the UCS; Everson

This proposal is in item 4 in the US document N[3045](#) and item 2 in Irish document N[3078](#).

Mr. Michael Everson: In year 2000, six characters were proposed in document N[2241](#). Four of the six (top two and bottom two) were considered by UTC and are included in the US contribution - the Alef and Ain characters have been proposed. The Yod is not included.

Disposition: Accept four characters as proposed in US contribution. See relevant resolution M48.24 below.

*Relevant resolution:*

*M48.24 (4 Latin Egyptological characters):*

*Unanimous*

With reference to document N2241 and item 4 in document N3045, WG 2 accepts the following four (4) characters to the Latin Extended-D block:

A722 LATIN CAPITAL LETTER EGYPTOLOGICAL ALEF  
A723 LATIN SMALL LETTER EGYPTOLOGICAL ALEF  
A724 LATIN CAPITAL LETTER EGYPTOLOGICAL AIN  
A725 LATIN SMALL LETTER EGYPTOLOGICAL AIN  
with their glyphs that are similar to those shown in document N2241.

### 8.27.1.2 Lao character name annotations

Input document:

[3137](#) Notes on Lao Character Names; Andrew West; 2005-10-30

This is item 6 in the US document [N3045](#) and item 4 in the UK document [N3065](#).

"The US is in favor of annotating the following Lao characters to mitigate partially their incorrect names which cannot be changed per policy of the standard itself. The supporting document is L2/05-357 (WG2 document number [N3137](#)):

U+0E9D LAO LETTER FO TAM (fo fon)  
U+0E9F LAO LETTER FO SUNG (fo fay)  
U+0EA3 LAO LETTER LO LING (ro)  
U+0EA5 LAO LETTER LO LOOT (lo) "

Discussion:

- a. Mr. Michel Suignard: The annotations arise because the user community usage of names is different from those published in the standard.
- b. Mr. Yasuhiro Anan: I want to ensure that input from the user community is sought.
- c. Dr. Ken Whistler: People who knew the language in the community had responded and identified the problem.

Disposition: Accept the proposed annotations for four Lao characters as in document [N3045](#). See relevant resolution M48.36 below.

*Relevant resolution:*

*M48.36 (Annotations for 4 Lao characters):*

*Unanimous*

With reference to item 6 in document N3045 WG 2 accepts to add annotations for the following four (4) characters in the Lao block:

0E9D LAO LETTER FO TAM (fo fon)  
0E9F LAO LETTER FO SUNG (fo fay)  
0EA3 LAO LETTER LO LING (ro)  
0EA5 LAO LETTER LO LOOT (lo).

### 8.27.2 Proposals from UK national body

Input documents:

[3065](#) Additions to AMD 3 - Tibetan, Mongolian, Myanmar, and Lao; UK NB - Andrew West; 2006-03-27

The following four items in document [N3065](#) have been covered under other agenda items.

- item 1. Four Tibetan characters from documents [N3011](#), [N3032](#) and [N3033](#) - see sections 8.1.2 and 8.1.3 on page 15.
- item 2. New Mongolian character from document [N3041](#) - - see section 8.15 on page 23.
- item 3. Seven Myanmar characters from documents [N3043](#) and [N3061](#) - see section 8.16.1 on page 24.
- item 4. New annotations for four Lao characters as per document L2/05-357 ([N3137](#)) - see section 8.27.1.2 on page 33.

### 8.27.3 Proposals from Irish national body

Input document:

[3078](#) Proposed additions to 10646 Amendment 3; Irish NB; 2006-04-06

The following two items in document [N3078](#) are covered by other agenda items.

- item 1. Seven new Myanmar characters in documents [N3043](#) and [N3061](#) - see section 8.16.1 on page 24.



item 2. Four new Latin characters for Egyptian transliteration from document [N2241](#) - see section 8.27.1.1 on page 32.

## 9 New contributions/feedback resulting from action items

### 9.1 Clarification - Hangul Syllables

Input documents:

<a href="#">2994</a>	Request to clarify text on conjoining Jamo - clause 26.1; Unicode Consortium and INCITS/L2 (US); 2005-09-14
<a href="#">2996</a>	Clause 26.1 - 10646: 2003; Kyongsok KIM, Dae Hayuk Ahn; 2005-09-15
<a href="#">3006</a>	Clarification concerning Hangul Syllables - 26.1; Project Editor - Michel Suignard; 2005-09-15
<a href="#">3045</a>	Additions to ISO/IEC 10646:2003; US National Body; 2006-02-22
<a href="#">3095</a>	Proposed modification to Clause 26.1 of ISO/IEC 10646:2003; NB Korea; 2006-04-24

Michel Suignard: Document [N3045](#) item 1 proposes new text towards 'Clarification concerning Hangul syllable method'. The change is really 'one' replaced by 'one or more'. A note re: AC00 to D7A3 is proposed. The proposed text is to address the concern that was raised in Sophia Antipolis meeting M47. It is not changing any of the normative parts. This text will also make the synchronization with Unicode 5.0 better. WG 2 should consider accepting this into current Amendment or a future one.

Discussion:

- a. Mr. Mike Ksar: Would like to know if the proposal in USA document is accepted in principle or not.
- b. Mr. Michel Suignard: As I said this is really a clarification of what is in current standard.
- c. Prof. Kyongsok Kim: We do not agree with it.
- d. Mr. Mike Ksar: Why did you not come up with your contribution earlier?
- e. Prof. Kyongsok Kim: Our position is the same as what we expressed at the last meeting. Precomposed syllables should not be combined with xxx code positions. Korea has adopted ISO/IEC 10646 and has specified that these combinations are not permitted in an Annex.
- f. Mr. Michel Suignard: If that is what you have adopted, you are not in compliance with ISO/IEC 10646.
- g. Prof. Kyongsok Kim: As you mentioned at the last meeting there is some inconsistency. As the standard is written, no one can/cannot comply with the standard -- because it is vague.
- h. Mr. Ienup Sung: I did not attend the last meeting. My information is only from documentation that is available. As long as clarification is made that should have been acceptable to Korea as well.
- i. Mr. Mike Ksar: Is this clarification good enough?
- j. Mr. Ienup Sung: Perhaps some additional clarification could be made such as 'does not allow combining precomposed Hangul syllables and Jamo components'; this may satisfy the Korean national body.
- k. Dr. Asmus Freytag: I cannot understand what is being said - perhaps a smaller group can get together to sort this out.
- l. Mr. Michel Suignard: Clause 25 in Amendment (used to be clause 24) - normalization forms -- defines behaviour for Korean characters. We are trying to explain consequences of normalization regarding Korean characters -- we are not trying to change anything else in the standard. I want to make sure everyone understands that clearly.
- m. Mr. Mike Ksar: I also want to ensure everyone understands that.
- n. Mr. Michael Everson: The ad hoc should go through the proposed text and get it sorted out if there are problems with the text.
- o. Dr. Asmus Freytag: Normalization is one of the areas in the standard that either we understand or not -- it cannot be changed due to the stability requirements. User communities rely on this being stable. Looks like national bodies have not understood that constraint that normalization is fixed. The aim is to eliminate confusion in the text and we can fix areas that could be fixed.
- p. Mr. Mike Ksar: If the Korean contribution does not address the normalization related items

(An ad hoc group of interested WG 2 experts met.)

- q. Prof. Kyongsok Kim: There will be need for proposing new characters. The change of text will be related to the text that is proposed to be added to the Amendment.
- r. Mr. Michel Suignard: Clarification text will be going into the Amendment.

- s. Mr. Mike Ksar: We could add the text to the Amendment now and we could propose additions later.
- t. Mr. Yasuhiro Anan: We would like to see the text clarifying what is going to be done.
- u. Mr. Michel Suignard: The proposed text is just clarifying what is in the standard today. Any addition to make processing of Old Hangeul easier is a separate entity -- that may involve additional text that goes with it.
- v. Prof. Kyongsok Kim: That is not my understanding. We will meet separately and resolve it.
- w. Mr. Mike Ksar: If there is no agreement then the text proposed by the editor must go in as clarification of current text in the standard.
- x. Dr. Asmus Freytag: We need as a minimum to put in clarification text to remove the contradiction in the current standard. It will be in the Amendment that will go out for ballot anyway.
- y. Mr. Mike Ksar: We talked about it in Sophia Antipolis and we already have a 6 months delay.
- z. Mr. Michel Suignard: there is some confusion on how to represent old Hangeul in the standard. Till we decide either to use sequence or precomposed values there is instability in the standard. In clause 26.1 we will remove one word, which occurs three times, and at the same time ask Korea to propose characters for precomposed Hangul. If the Korean proposal does not come then we will be stuck with problem of no precomposed Jamos for Old Hangul. We have removed the term 'characters' which occurs three times.
- aa. Mr. Mike Ksar: The word changes will go in current Amendment. We would like to see the Korean proposal for the precomposed Jamos model for discussion much earlier than in September.
- bb. Dr. Asmus Freytag: UTC has a meeting in August and would like to have it before then.
- cc. Mr. Mike Ksar: We should have the proposal by 15 July for posting to WG 2 site.
- dd. Mr. Michel Suignard: The proposal breaks the model that we have to some extent and national bodies should have opportunity to study.

Disposition: Accept the ad hoc recommendation. We will remove 'character' that occurs three times in clause 26.1 in Amendment 3. Korea is to propose a New Model for using Hangeul Jamos by 15 July 2006. See relevant resolutions M48.30 and M48.31 below.

*Relevant resolutions:*

*M48.30 (Hangul clarification):* *Unanimous*  
 WG 2 accepts removing the word 'character' (that occurs three times in the parenthetical notations in the first paragraph) in clause 26.1 in the standard.

*M48.31 (Model for Hangul Jamos):* *Unanimous*  
 Republic of Korea is invited to prepare a document on 'a new model for using Hangul Jamos' and send to the convener no later than 2006-07-15, for distribution to and review by WG 2 experts.

## **9.2 Improving formal definition for control characters**

Input document:

[3046](#) Improving formal definition for control characters; Michel Suignard, Project Editor; 2006-02-22

Mr. Michel Suignard: ITU in its liaison statement (see documents [N3013](#) and [N3026](#) and related discussion in section 5.1 on page 12) is asking for changes to control characters. They want names for the control characters. They are complaining that ISO/IEC 10646 does not explicitly describe the control characters or normatively point to ISO/IEC 6429. We already do it. Looks like ITU has not read the standard. Even in ISO/IEC 6429 it is not easy to discover what the names of the control characters are. Suggestion is to include an extract of the control characters from ISO/IEC 6429 by putting meaningful names for these characters. Document [N3046](#) is the first draft in an attempt at it.

Discussion:

- a. Dr. Umamaheswaran: ISO/IEC 10646 has also escape sequences that permits alternate use of Controls in C0 and C1. We can identify that as default in the absence of such escapes, the ISO/IEC 6429 default is used. It will be useful to have the information within ISO/IEC 10646 as service to the users.
- b. Dr. Asmus Freytag: We do not have the escape mechanism like in ISO/IEC 10646. In view of synchronization we could entertain the proposal. For documenting the default, in Unicode, we have given aliases for each of the controls copying from ISO/IEC 6429. If ISO/IEC 10646 wants to document this in an informative way we could use the information from Unicode. If



ISO/IEC 10646 wants to make it a normative one then we may have a synchronization issue. The aliases may not be unique. It is different from the proposed attempt to regularize the controls.

- c. Mr. Michael Everson: Neither ISO/IEC 10646 nor Unicode standard have encoded these as characters -- and they have not assigned character names. I would not object to tidying up this area. For empty control codes we could leave them alone.
- d. Dr. Ken Whistler: It will be simpler to satisfy most people to give these characters names -- for those using ISO/IEC 6429 as default there will be no issues. For others who would swap out they can overwrite these. It could give the impression that ISO/IEC 10646 is trying to fix these. One way to overcome that concern is that is to name these as 'control character defaulted to xxxx in ISO/IEC 6429'. Even though it is named as xxxx in the standard, one has to look into ISO/IEC 6429 to understand that these can be changed by protocols to others. The group that wants to swap these will come to us.
- e. Dr. Asmus Freytag: We are essentially treading too lightly on the Unicode perspective that treats controls as characters -- we have given them pseudo names, and we cannot change these from stability point of view that will cause a synchronization problem. Aliases allow us to call them what we want - copying for example ISO/IEC 6429.
- f. Mr. Michel Suignard: ITU is asking for names.
- g. Dr. Asmus Freytag: The confusion arises because we have a vacuum.
- h. Dr. Umamaheswaran: The cleanest way is to copy what we have in the Unicode standard. Allows us to guarantee synchronization and stability.
- i. Dr. Ken Whistler: Agree with the general approach outlined by Dr. Umamaheswaran. I also agree with Dr. Asmus Freytag in that we have not addressed the control functions in the standard. Many of the suggestion made by Mr. Michel Suignard are valuable. Simply by adding the names would solve the general problem. To address ITU concerns, we could add an informative annex - we could copy the ISO/IEC 6429 default into that informative annex.
- j. Mr. Dae Hayuk Ahn: Why are some names short like DEL, IND and others are expanded.

Disposition: Reference N[3046](#) and ITU-T SG17 Request item a. Agree to add information regarding control sets C0 and C1 in the standard; Add normative definitions for missing terms. Add informative annex showing content of C0 and C1 sets with 'annotations' in 'lower case' of the form (control) (long name from ISO/IEC 6429). Based on document N[3046](#) and discussion at this meeting the editor is to prepare the appropriate text. See relevant resolution M48.33 below.

*Relevant resolution:*

*M48.33 (Information on control characters):*

*Unanimous*

With reference to document N3046 and to item a regarding control characters in ITU-T SG17 request document N3013, WG 2 resolves to add to the standard:

- a. normative definitions for relevant terms related to control characters and control functions
- b. an informative annex showing the default content of 0000 to 001F, 007F, and 0080 to 009F blocks and charts similar to those published in the Unicode Standard, and names in the form of annotations - "(control) (long name from ISO/IEC 6429)".

WG 2 further instructs its project editor to prepare appropriate text including the feedback from the discussion on document N3059 at meeting 48.

### **9.3 Ideographic Variation Database**

Input document:

[3035](#) Ideographic Variation Database; USNB - Eric Muller; 2006-02-02

Dr. Ken Whistler: First part of this document is to notify WG 2 of Unicode publication of UTS37 which is the registration of Ideographic Variation sequences -- based on vendor requirements to identify large collections of variation sequences of Han characters. It gives a way to specify the specific variation in a specific collection. This is using the Variation Selection character -- Han Character + VS. This raises the synchronization issue with ISO/IEC 10646. To date we are in synch using the VS-s -- allowing only Variation sequences in the standard as allowed. In order to allow use of UTS 37 as registration means for Ideographic VS-s the US national body is asking for a change of text in the standard to point to UTS 37 and allowing for VS-s that are not explicitly listed in the standard. Document N[3035](#) has specific proposed text. The sooner we include such text in the standard, the better. Specifically the request is to change clause 20.4 in Amendment 3.

Discussion:

- a. Mr. Mike Ksar: Is IRG aware of this?
- b. Mr. Yasuhiro Anan: US expert have attended the IRG and explained this to IRG. So IRG is aware. IRG has no position on this. I am a little concerned about the registration fees as far as IRG's work is concerned. If IRG wantx to make use of this registration database the IRG has to find some funds.
- c. Dr. Ken Whistler: UTS 37 has a mechanism for registration -- that indicates a nominal registration fee to keep some control on requests -- but that fee can be waived if there is a reason to waive that.
- d. Dr. Asmus Freytag: The fee mechanism was there to control registration only by serious requesters versus an open registry. The fee MAY be charged.
- e. Mr. Yasuhiro Anan: In terms of IRG work there will not be a fee.
- f. Mr. Mike Ksar: The Unicode Consortium can be the registrar on behalf of IRG.

Disposition: Accept proposed text in document N[3035](#). See relevant resolution M48.32 below.

*Relevant resolution:*

*M48.32 (Ideographic Variation Database):*

*Unanimous*

WG 2 accepts the proposed replacement of two paragraphs in clause 20.4 with three new paragraphs and one normative reference to UTS #37, as detailed in document N3035, in order to include by reference the Variation Sequences defined in the Ideographic Variation Database in that standard.

#### **9.4 Collection identifiers for Japanese subsets**

Input document:

[3091](#), [3091-A](#)

Proposal for collections ; Japan; 2006-04-20

Mr. Yasuhiro Anan: Document N[3091](#) requests for five new collections to identify ISO/IEC 10646 subsets used in Japanese national standards. We refer to non-ISO/IEC 10646 standards such as JIS X0213 etc. Vendors have started implementing JIS X0213 repertoire in ISO/IEC 10646 encoding. The vendor intention is to support the repertoire of the legacy encoding standards using ISO/IEC 10646. This causes confusion. To avoid the confusion the Japanese national bodies has recognized the need to identify the repertoires of the existing national standards as ISO/IEC 10646 collections. We reduced the number of sets from 7 to 5 to reduce the work of the editor. Currently because USIs cannot be included, the parts needed to create the USIs are included in the collection. The collections would allow us to represent extensions.

Mr. Yasuhiro Anan described the basis for each of the five proposed collections 1 to 5 in document N[3091](#). He explained the use of these collections -- based on the two core sets BASIC and COMMON Japanese and extensions thereof.

Discussion:

- a. Mr. Mike Ksar: the collections following the current procedures are in a separate .zip file identified as N[3091-A](#).
- b. Dr. Ken Whistler: Why is it NOT sufficient to keep these collections only in JIS? If these are included in ISO/IEC 10646, what would Japanese national body do when one of the JISs is revised or extended? Why can't Japan define another JIS which defines just the collections required in Japan?
- c. Mr. Yasuhiro Anan: We are currently updating JIS X0221 - corresponding to ISO/IEC 10646: 2003. We have to update the descriptions of the original annexes. Without the formal definitions in ISO/IEC 10646 the national standard will not be identical to the ISO standard. The National body discussed the second question -- and users would prefer to find all the information in one standard.
- d. Dr. Ken Whistler: The intent of Japanese national body is to include these in ISO/IEC 10646 and when they get translated into Japanese the users of the Japanese standard will have one standard to use.
- e. Mr. Yasuhiro Anan: We do not plan to update the legacy sets; and when we want any expansions we will request WG 2.
- f. Mr. Mike Ksar: Do we have any concerns?
- g. Mr. Michel Suignard: It would take some work to present the information in a usable manner.
- h. Mr. Yasuhiro Anan: Can we expect Amendment 3 to be available by 2007?

- i. Dr. Umamaheswaran: If we don't have any show stoppers we should include these in Amendment 3.
- j. Mr. Erkki Kolehmainen: With the extended collection definition we accepted, can Japan use these?
- k. Dr. Umamaheswaran: Mr. Yasuhiro Anan has told us that by identifying only the components in the sequences they can keep the sequences open.

Disposition: Accept five collections as defined in document [N3091-A](#) for inclusion in Amendment 3.. Editor and Japan (Mr. Yasuhiro Anan) will work together to get the right format for inclusion in the standard. See relevant resolution M48.29 below.

*Relevant resolution:*

*M48.29 (5 collections corresponding to Japanese standards):*

*Unanimous*

With reference to the five requested collections in document N3091, WG 2 accepts to define five new collections in the standard based on information provided in attachment (N3091-A) to document N3091.

## 10 Amendment 3

### 10.1 Proposed disposition of comments on Amendment 3

Input document:

[3071](#) Summary of Voting/Table of Replies - PDAM 3; SC2 Secretariat, 02n3852; 2006-04-07

[3049](#) Proposed Disposition of Comments PDAM 3; Michel Suignard; 2006-04-14

Output documents:

[3098](#) Disposition of Comments PDAM 3; Michel Suignard; 2006-04-26

[3059](#) Repertoire Charts – Amendment 3 – 1004 characters; Asmus Freytag; 2006-04-27

Mr. Michel Suignard: Comments were received from Ireland, Japan, United Kingdom, and USA. Document [N3049](#) is the disposition of those comments. The disposition is organized per country. Note – The full content of the ballot comments (minus some character glyphs) have been included in this document to facilitate the reading. The dispositions are inserted in between these comments and are marked in **Underlined Bold Serif text**, with explanatory text in *italicized serif*.

#### 10.1.1 Ireland: Disapproval

##### T1. Page 10 - Row 1C: Lepcha.

Mr. Michael Everson: Recalling a WG 2 request clarification on two Lepcha characters - document [N2947R](#) has the following .. "NYIN-DO is used only when there is no vowel matra; KANG is used in all other instances. .... The two are encoded separately because as named characters they may be required in isolation in discussions of the writing system; positional variants would not support this efficiently." At meeting M47 clarification was requested as to why they cannot be glyph variants of each other.

Ireland has provided reference Mainwaring & Grünwedel's *Dictionary of the Lepcha Language*, pp 10 and 106 and an explanation as to why these should be encoded separately (as is in current amendment). NYIN-DO is at 1C34 and KANG is at 1C35. Whether or not the font can determine the shape / character etc. was one of the consideration. If context is used for determination of which of the two characters, then they cannot be displayed in isolation when needed. The name of the character itself is related to other characters in the standard. Original contribution was based on an implementation, and was unified based on that. Ireland thinks that these should not be unified.

Discussion:

- a. Mr. Michel Suignard: US ballot comment T.1 also requests this clarification. The US comment seems to be accommodated by the Irish comment. If we accept the Irish comment does the US have any objection.
- b. Dr. Ken Whistler: US will go along with Irish comment.

Disposition: Irish T.1 is noted. US comment T.1 is accepted. There is NO change in Amendment as a result.

##### T2. Page 14 - Row 2C: Latin Extended-C.

Mr. Michel Suignard: There are two identical comments - one from Ireland and one from UK. Both request to remove Mayanist letters 2C6F and 2C70 and replace them with 8 new characters.

Disposition: See discussion and relevant resolution M48.2 under agenda item 8.11 on page 20 on this topic.

##### T.3. Pages 18-21 - Rows A5-A6: Vai script.

Ireland requests the insertion of thirteen new characters to the Vai character set.

Further justification for these additions is provided in document N3081R.

Disposition: See discussion and relevant resolution M48.3 under agenda item 8.24 on page 31.

#### **T.4. Page 24 - Row A8: Saurashtra.**

Name of U+A8B4 to be changed from SAURASHTRA LETTER UPAKSHARA to SAURASHTRA LETTER HAARU or SAURASHTRA CONSONANT SIGN HAARU.

Disposition: See discussion and relevant resolution M48.4 under agenda item 8.17.1 on page 25.

#### **E.1 Page 4 - Row 03: Greek and Coptic.**

Glyph for 037E GREEK QUESTION MARK is now incorrect. The complete chart on Greek should be reviewed for glyphs.

Disposition: Noted. However, this character was not one of those accepted in Amendment 3. The character shape in the charts has to be corrected for the next version of Amendment 3.

#### **E.2 Page 18 - Row A5: Vai.**

The glyph for A553 (new A555) VAI SYLLABLE KPAN is incorrect. It should have an x-shape, not a +-shape.

Disposition: Accept. New shape is in N3059.

Mr. Michael Everson: Ireland reverses its vote to Approval.

#### **10.1.2 Japan, Yes with comments**

J1 The title of the Amendment should read "Lepcha, Oi Chiki, Saurashtra, Vai and other characters" (i.e., no comma before "and".)

Disposition: Not accepted - current title is in line with style used.

Comments J2, J3 were accepted.

#### **10.1.3 United Kingdom: Yes with comments:**

##### **T1. Page 14 Table 103 : LATIN LETTER TRESILLO and LATIN LETTER CUATRILLO**

Disposition: Same comment as Irish T.2 on Mayanist letters. See discussion and relevant resolution M48.2 under agenda item 8.11 on page 20 on this topic.

#### **10.1.4 USA: Yes with comments**

##### **T.1 Lepcha character needing clarification**

1C35 LEPCHA CONSONANT SIGN KANG requires further explanation

Disposition: Accepted; no change is needed based on input under Irish comment T.1 above.

#### **10.1.5 Final disposition**

The editor will prepare the final disposition of comments document incorporating the discussion at the meeting. The following is the relevant resolution M48.5.

*Relevant resolution:*

*M48.5 (PDAM3 Disposition of Comments): Unanimous*

WG 2 accepts the disposition of comments on PDAM3 ballot in document N3098 and instructs its editor to prepare the final text of Amendment 3 incorporating the dispositions. The following are noted in particular:

- a. Changes noted in resolution on 'changes to Mayanist characters'
- b. Changes noted in resolution on 'changes to Vai script'
- c. Correct the shape of new A555 VAI SYLLABLE KPAN
- d. Change noted in resolution on 'change name of one Saurashtra character'

#### **10.2 Progression of Amendment 3**

Mr. Mike Ksar: There are two items that could require another ballot for Amendment 3. One is the synchronization with Unicode standard. The other item is if we include major technical changes such as inclusion of Myanmar changes to Amendment 3 it would require another PDAM ballot. The potential delay about 3 months would be the result. I would need feedback from the national bodies on this item. The future Amendment and current Amendment 3 could be combined into a single Amendment if we decide to go for another PDAM ballot for Amendment 3. We also don't want to face urgent customer requests etc. that we faced in Sophia Antipolis.

Discussion:

- a. Dr. Umamaheswaran: If we delay Amendment 3, we will be out of synch with Unicode 5.0 with a few characters that are in V5.0 but not in Amendment 2. Also the Ideographic Variation Database we just accepted will be out of synch.

- b. Dr. Asmus Freytag: Schedule for V5.0 book is behind our expectation. It is synchronized with Amendment 2. Until Unicode 5.0 work is completed, Unicode also has lack of resources to synchronize with any new Amendments of ISO/IEC 10646. The community has come to anticipate synchronization with specific Amendment (s). We like the idea of making a single Amendment to synchronize with rather than having to synchronize with multiple Amendments.
- c. Mr. Mike Ksar: My proposal is to combine into a single Amendment -- merging Amendment 3 and future Amendment. I want to receive feedback from different national bodies.
- d. Mr. Michael Everson: From tracking point of view it is easier to keep track of a single Amendment under ballot than multiples. There is also the risk of failure due to lack of participation by national bodies.
- e. Mr. Erkki Kolehmainen: The delay in new characters will not be a major problem for us.
- f. Mr. U Thein Oo: There is agreement in principle to have two technical ballots.
- g. Prof. Kyongsok Kim: It will be OK to deal with a single Amendment.
- h. Mr. Yasuhiro Anan: We are in favour of two ballots on Amendment 3 due to additional technical changes.
- i. Ms. Toshiko Kimura: It would be the natural thing to do if we have major technical changes.

Disposition: Accept processing of Amendment 3 as second PDAM3, with revised schedule. Include all items accepted at this meeting into this single Amendment. See relevant resolution M48.37 below.

*Relevant resolution:*

*M48.37 (Progression of Amendment 3):*

*Unanimous*

WG 2 resolves to include all the items accepted for inclusion in the standard noted in the various resolutions above along with the changes arising out of the disposition of comments into Amendment 3. WG 2 instructs its project editor to prepare the revised text of Amendment 3 and forward it along with the disposition of comments document N3098 to the SC 2 secretariat for a second PDAM ballot. The final set of charts and names lists are in document N3059. The schedule for this work item is revised accordingly to: PDAM 2006-05-31, FPDAM 2006-10-31, and FDAM 2007-05 (starting dates are given).

## 11 Publication issues

There were none identified for this meeting.

## 12 IRG status and reports

Input documents:

[3062](#) IRG 25 Resolutions; Lu Qin - IRG; 2005-12-02

[3064](#) Summary Report IRG 25; Lu Qin - IRG; 2006-04-11

[02N3814](#) Table of Replies for ISO/IEC 10646/FDAM1, Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- Amendment 1: Coptic, Georgian and other characters ITTF ISO/IEC 10646/FDAM1 has been approved for publication. The comments received are forwarded to the project editor. Summary of Voting/Table of Replies; 2005-09-20

### **Item 1: Future meetings.**

Dr. Lu Qin: M26 is scheduled for Hue, Vietnam, 2006-05-5/9. Other meetings IRG 27 in Taipei 2006-11-27/12-01. Others are listed in document [N3064](#). USA is to confirm hosting IRG 29.

Disposition: Approve the meetings. See relevant resolution M48.42 on page 43 in section 15.2.

### **Item 2: Ext C1 Quality**

Dr. Lu Qin: In C1 editing work we are encountering quality issues. The repertoire of CJK\_C1 could not be delivered to WG 2 at this meeting. We will be using the Ideographic Description Sequences to describe each Han proposed. Using that IDS database we will be able eliminate duplicates etc. IRG will also maintain a CJK Error List. A protocol is to be established and error list could be sent to WG 2 on request.

### **Item 3 - Ext C1 split**

Dr. Lu Qin: We have decided to provide only a subset of about 5000 (4629 characters) which are considered to be STABLE -- i.e. there is no comment, no controversy, and they have been examined at least 3 times. These are to be confirmed at the IRG meeting. The remaining C1 characters are moved into CJK C2. We will submit C1 only if the quality is OK and IRG is comfortable with it.

Discussion:



- a. Mr. Mike Ksar: If we come up with C1 in Japan meeting and if we have to delay Amendment 3 we will have problem including these in Amendment 3. We may have to open up another Amendment beyond for C1.
- b. Dr. Umamaheswaran: If C1 is input at this time to Amendment 3 we could include it and let the ballot review take care of any instabilities.
- c. Mr. Michel Suignard: We should have the production capabilities -- i.e. fonts. All the parties interested in IRG work we could agree on it.
- d. Dr. Lu Qin: I can report to you right away after the next IRG meeting. Convener can send it in for review.

#### **Item 4 CJK Strokes**

This topic is discussed under agenda item 8.19 on page 28.

#### **Item 5 - Old Hanzi**

Dr. Lu Qin: We will be working on Oracle Bone inscription part of Old Hanzi first. All the information will be in soft copy XML formats. The working group has not submitted any dates yet. Other IRG members will be submitting their information.

Discussion:

- a. Dr. Ken Whistler: Do we know how long it will take to identify all the Oracle Bone inscriptions?
- b. Dr. Lu Qin: No idea at this time.
- c. Mr. Michael Everson: Should we update the roadmap?
- d. Dr. Lu Qin: Extension C1 is about 5000.

#### **Other items:**

Mr. Michel Suignard: In the FDAM ballot of Amendment 1 - a comment was made by Korea (and others) (see document [SC 2 N3814](#) -- several comments were about issues in the Source reference. I would request IRG to address those concerns in that document. There are duplicate source entries and these need to be fixed. Also some IICORE (informative) issues, and G-Source issues .. For the sake of being clean I request IRG to review document N3814.

Discussion:

- a. Mr. Mike Ksar: Why are NBs making technical comments at the FDAM stage?
- b. Ms Toshiko Kimura: Comments can be made by any national body. Only editorial changes are normally expected.

Disposition: IRG is to address the CJK-related ballot comments in FDAM1 in document [SC 2 N3814](#) regarding . WG 2 acknowledges the rationale for delay in C1 work and accepts the delay in C1 submission. WG 2 endorses IRG rapporteur for another term. See relevant resolutions M48.39, M48.40 and M48.41 below.

#### *Relevant resolutions:*

*M48.39 (CJK Ext C1):*

*Unanimous*

WG 2 takes note of item 3 in document N3064, in particular that CJK Ext C1 repertoire is expected to be finalized at IRG meeting 26 in June 2006, and accepts the delay in its submission for WG 2 consideration at its next meeting.

*M48.40 (CJK comments on FDAM1):*

*Unanimous*

WG 2 instructs the IRG to review and prepare a response to the various issues raised regarding CJK Unified Ideographs in the FDAM1 ballot response in document SC 2 N3814.

*M48.41 (IRG rapporteur):*

*By Acclamation*

WG 2 nominates Dr. Lu Qin as the IRG Rapporteur for one more term, and requests SC 2 to approve this appointment.

### **13 Defect reports**

There were none reported at this meeting.

### **14 Liaison reports**

#### **14.1 The Unicode Consortium**

Input documents:

[3016](#) Character Count Chart; Ken Whistler; 2005-11-21

[3057](#) JTC1 ARO - SC2 on Unicode; Mike Ksar, SC2 Secretariat 02n3841; 2006-03-16

Dr. Asmus Freytag: Document [N3016](#) contains the character counts in Unicode. We have 99020 now (Amendment 2 has 4 less). Amendment 3 would add 1004 new characters.



Discussion:

- a. Mr. Michel Suignard: A fix up for collection to reflect up to end of Amendment 2 has to be looked at.
- b. Dr. Umamaheswaran: The more useful collection id would be for the repertoire of Unicode 5.0 that include the four characters that are in V5.0 but not in Amendment 2.

Disposition: Update the collection definition for collection 307 to correspond to V5.0 repertoire -- i.e. to include four Devanagari characters from Amendment 3 - .097B, 097C, 097E, 097F (that are not in the standard up to Amendment 2) See relevant resolution M48.35 below.

*Relevant resolution:*

*M48.35 (Updating collection 307 definition):*

*Unanimous*

WG 2 accepts updating definition of collection '307 Unicode 5.0' to include the four Devanagari characters 097B, 097C, 097E and 097F (encoded in Amendment 3) to correspond to the repertoire of Unicode 5.0.

Mr. Mike Ksar: Document N[3057](#) - is the ARO document that was sent to SC 2.

Discussion:

- a. Dr. Umamaheswaran: Is there any statement about Trademarks etc. in the ARO document that we can reference in the response to ITU-T.
- b. Mr. Tatsuo Kobayashi: We can point to this ARO document and link to the trademark section, as well as the JTC 1 directives Annex N.

Disposition: Include the reference and request to forward document N[3057](#) to ITU-T as part of liaison response to item f in ITU-T document. See relevant resolution M48.38 on page 12 under section 5.1.

## **14.2 IETF**

### **14.3 JTC 1/SC 22**

Input document:

[3008](#) SC2 to SC22 Liaison Report; Mike Ksar (SC2N3808); 2005-09-15

The document contains the liaison report prepared by Mr. Mike Ksar to SC 22 from SC 2. It is for information of the delegates. There was no discussion.

### **14.4 W3C**

#### **14.5 UC Berkeley - request for liaison**

Input document:

[3068](#) Request for Liaison from UC Berkeley; UC Berkeley; 2006-04-03

Ms. Deborah Anderson described the document N[3068](#).

Mr. Mike Ksar: We welcome UC Berkeley and all the contributions you have made to work of WG 2. It will be a valuable addition. SC 2 will discuss this item and will establish the liaison formally.

Disposition: Support the request from UC Berkeley and request SC 2 to approve the request in N[3068](#). See the relevant resolution M48.44 below.

*Relevant resolution:*

*M48.44 (Liaison request from UC Berkeley):*

*Unanimous*

WG 2 requests SC 2 to accept the liaison request from UC Berkeley in document N3068.

## **15 Other business**

### **15.1 Web Site**

Mr. Mike Ksar: The WG 2 web site will be down during maintenance periods. I have updated based on some suggestions received. See the relevant resolution of appreciation M48.45 below.

*Relevant resolution:*

*M48.45 (Appreciation to DKUUG for web site support):*

*By Acclamation*

WG 2 thanks DKUUG and its staff, in particular Mr. Kristen Nielsen, for its continued support of the web site for WG 2 document distribution and the e-mail server.

## 15.2 Future meetings

Input documents:

- [3030](#) Invitation from German NB to host SC2WG2 meeting 50 in Frankfurt; German NB; 2006-01-18
- [3067](#) Meet 49 Invite by Japan; ITSCJ; 2006-04-06
- [3093](#) Logistical information for the ISO/IEC JTC1/SC2/WG2 meeting 49 in Akihabara, Tokyo - Japan; Japan; 2004-04-24

- Meeting 49 – 2006-09-25/29 (5 days) - Tokyo, Japan  
Mr. Mike Ksar: Document [N3067](#) is the formal invitation from Japan. Document [N3093](#) has all the logistics.  
I will give the floor to Japan. The meeting will be for 4 and half days. There is no meeting or facilities fee.  
Mr. Kazuhito Ohmaki: The next WG 2 meeting will be in Akihabara section of Japan. The meeting is in a new building 11th floor of a 30 floor building. Information about nearby hotels, transportation etc. are given in document [N3093](#). Registration form and hotel lists and rates, map of hotel locations are also in that document.
- Meeting 50 – April 2007, Frankfurt, Germany  
Mr. Mike Ksar: Germany is our host. Frankfurt is the location. See the invitation in document [N3030](#). They gave us two dates to choose from -- week of the 16th or week of 23rd April 2007. We should take the week of 16 - 20 of April (4.5 days). If there are changes I will let everyone know.
- Meeting 51 – Fall 2007 – seeking host.  
Dr. Lu Qin: Hong Kong is likely site -- I will check and let you know.  
We have at least one backup host. Japan is willing to be backup.  
Ms. Wei Lin-Mei: I will check about hosting in Taiwan.
- Meeting 52 – April 2008 – Republic of Korea.  
Prof. Kyongsok Kim, Mr. Dae Hayuk Ahn: Republic of Korea is willing to host. Will confirm.

See the relevant resolution M48.42 below.

*Relevant resolution:*

*M48.42 (Future meetings):*

*Unanimous*

WG 2 meetings:

- Meeting 49 – 2006-09-25/29, Tokyo, Japan
- Meeting 50 – 2007-04-16/20, Frankfurt, Germany
- Meeting 51 - Fall 2007 - Japan or Hong Kong or Taipei to be confirmed)
- Meeting 52 - April 2008, Republic of Korea (to be confirmed)

IRG meeting:

- IRG #26 - 2006-06-5/9, Hue, Vietnam
- IRG #27 - 2006-11-27/12-01, Taipei (host: TCA)
- IRG #28 - 2007 May / June, Hangzhou, China (to be confirmed)
- IRG#29 - 2007 November, San Jose, California, USA (host Adobe).

## 15.3 Closing

### 15.3.1 Approval of resolutions of meeting 48

Output document:

- [3014](#) Towards an encoding of the Kaithi script in the SMP of the UCS; Michael Everson - Individual Contribution; 2005-11-06

1004 characters have been accepted for inclusion in Amendment 3 balloting as PDAM3-2 (adjusted from previous 517).

Experts from Canada, Finland, Ireland, Japan, Korea (Republic of), Poland, Taipei Computer Association (Liaison), the Unicode Consortium (Liaison) and the USA were present when the resolutions prepared by Dr. Umamaheswaran and the drafting committee were reviewed and adopted with some changes. The adopted resolutions are in document [N3014](#).

From a total of 99016 characters at the end of Amendment 2, another 517 characters were in PDAM3. After adjusting the content of Amendment 3 based on disposition of PDAM3 ballot comments, and adding characters from new contributions considered at this meeting the second PDAM3 will contain 1004 additional characters. This brings the total to 100020 characters to date.

## Appreciation

Relevant resolution:

M48.46 (Appreciation to Host):

By Acclamation

WG 2 thanks the US national body, Microsoft Corporation and its staff, and the Unicode Consortium and its staff, in particular Ms. Magda Danish, for hosting the meeting, their kind hospitality, and for providing excellent meeting facilities.

### 15.4 Adjournment

Meeting adjourned at 12:35h local time.

## 16 Action Items

All action items recorded in the minutes of the previous meetings from M25 to M46, have been either completed or dropped. Status of outstanding action items from earlier meeting M47, and new action items from the latest meeting M48 are listed in the tables that follow.

- Meeting 25, 1994-04-18/22, Antalya, Turkey (document N1033)
- Meeting 26, 1994-10-10/14, San Francisco, CA, USA (document N1117)
- Meeting 27, 1995-04-03/07, Geneva, Switzerland (document N1203)
- Meeting 28, 1995-06-22/26, Helsinki, Finland (document N1253)
- Meeting 29, 1995-11-06/10, Tokyo, Japan (document N1303)
- Meeting 30, 1996-04-22/26, Copenhagen, Denmark (document N1353)
- Meeting 31, 1996-08-12/16, Québec City, Canada (document N1453)
- Meeting 32, 1997-01-20/24, Singapore (document N1503)
- Meeting 33, 1997-06-30/07-04, Heraklion, Crete, Greece (document N1603)
- Meeting 34, 1998-03-16/20, Redmond, WA, USA (document N1703)
- Meeting 35, 1998-09-21/25, London, UK (document N1903)
- Meeting 36, 1999-03-09/15, Fukuoka, Japan (document N2003)
- Meeting 37, 1999-09-17/21, Copenhagen, Denmark (document N2103)
- Meeting 38, 2000-07-18/21, Beijing, China (document N2203)
- Meeting 39, 2000-10-08/11, Vouliagmeni, Athens, Greece (document N2253)
- Meeting 40, 2001-04-02/05, Mountain View, CA, USA (document N2353), and
- Meeting 41, 2001-10-15/18, Singapore (document 2403)
- Meeting 42, 2002-05-20/23, Dublin, Ireland (document N2453)
- Meeting 43, 2003-12-09/12, Tokyo, Japan (document N2553)
- Meeting 44, 2003-10-20/23, Mountain View, CA, USA (document N2653)
- Meeting 45, 2004-06-21/24, Markham, Ontario, Canada (document N2753)
- Meeting 46, 2005-01-24/28, Xiamen, China (document N2903)

### 16.1 Outstanding action items from meeting 47, 2005-09-12/15, Sophia Antipolis, France

Item	Assigned to / action (Reference resolutions in document N2954, and unconfirmed minutes in document N2953 for meeting 47 - with any corrections noted in section 3 of document N3103 from meeting 48).	Status
AI-47-5	<b>IRG Rapporteur (Dr. Lu Qin)</b> <i>To take note of and act upon the following items.</i>	
c.	With reference to discussion in meeting 47, regarding IICORE and safe characters for security, IRG is requested to review and feedback on UTS 36 for safe characters and to RFC 3743. M48 - in progress.	In progress.
AI-47-8	<b>Ireland (Mr. Michael Everson)</b>	
c.	To work with the author Mr. Nick Nicholas to provide supporting examples for Greek epigraphical characters proposed in document N2946, and accepted per resolution M47.23. M48 - in progress.	In progress.

### 16.2 New action items from meeting 48, 2006-04-24/27, Mountain View, CA, USA

Item	Assigned to / action (Reference resolutions in document N3104, and unconfirmed minutes in document N3103 for meeting 48 - this document you are reading)	Status
AI-48-1	<b>Meeting Secretary - Dr. V.S. UMaaheswaran</b>	
d.	To finalize the document N3104 containing the adopted meeting resolutions and send it to the convener as soon as possible.	Completed; see document N3104.
e.	To finalize the document N3103 containing the unconfirmed meeting minutes and send it to the convener as soon as possible.	Completed; see document N3103.
f.	To incorporate the clarification text regarding CJK Strokes in SIP vs SMP in the next	

Item	Assigned to / action (Reference resolutions in document N3104, and unconfirmed minutes in document N3103 for meeting 48 - this document you are reading)	Status
	snapshot of roadmap.	
<b>AI-48-2</b>	<b>Convener - Mr. Mike Ksar</b>	
e.	M47.37 (Roadmap snapshot): WG2 instructs its convener to post an updated snapshot of the roadmaps (document N3005) as soon as it is ready to the WG2 web site.	Completed; document N3005 is posted to WG2 site.
f.	Request SC2 to ask ISO/IEC 14651 editing group to take note of collation information regarding CJK Strokes in document N3063.	Completed; was done during editing meeting in Mountain View, CA.
g.	Request SC2 to circulate revised Principles and Procedures document N3002 (format changes) to SC2 member bodies.	Completed; SC2 has distributed.
h.	<p>To add the following carried forward scripts to next meeting agenda, if any updates are available:</p> <ul style="list-style-type: none"> <li>p. Manichaeian script – document N2544</li> <li>q. Avestan and Pahlavi – document N2556</li> <li>r. Dictionary Symbols – document N2655</li> <li>s. Invisible Letter – document N2822</li> <li>t. Palestinian Pointing characters - document N2838</li> <li>u. Babylonian Pointing characters – document N2759</li> <li>v. Bantu Phonetic Click characters – document N2790</li> <li>w. Samaritan Pointing - 12 Hebrew characters – document N2758</li> <li>x. Dominoe Tiles and other Game symbols – document N2760</li> <li>o. Four Tibetan charactes for Balti -- N3010</li> <li>p. Kaithi script -- N3014</li> <li>q. Karen, Shan and Kayah characters -- N3080</li> <li>y. Arabic Math proposals -- N3085, N3086, N3087, N3088 and N3089Cyrillic combining marks -- N3097</li> </ul>	
<b>AI-48-3</b>	<b>Editor of ISO/IEC 10646: Mr. Michel Suignard with assistance from contributing editors</b>	
	<i>To prepare the appropriate amendment texts, sub-division proposals, collection of editorial text for the next edition, corrigendum text, or entries in collections of characters for future coding, with assistance from other identified parties, in accordance with the following:</i>	Completed; see PDAM3-2 document N3109.
z.	<p>M48.2 (Changes to Mayanist characters): With reference to document N3082, and in response to PDAM3 ballot comments in document N3071 from Ireland, UK and USA, regarding Mayanist characters, WG2 accepts the following:</p> <ul style="list-style-type: none"> <li>a. Add two new characters to the Latin Extended-D block: <ul style="list-style-type: none"> <li>A726 LATIN CAPITAL LETTER HENG</li> <li>A727 LATIN SMALL LETTER HENG</li> </ul> with their glyphs that are similar to those shown in document N3082. </li> <li>b. Replace the following two (2) caseless characters in Latin Extended-C block: <ul style="list-style-type: none"> <li>2C6F LATIN LETTER TRESILLO</li> <li>2C70 LATIN LETTER CUATRILLO</li> </ul> with the following eight (8) case-paired characters in Latin Extended-D block: <ul style="list-style-type: none"> <li>A728 LATIN CAPITAL LETTER TZ</li> <li>A729 LATIN SMALL LETTER TZ</li> <li>A72A LATIN CAPITAL LETTER TRESILLO</li> <li>A72B LATIN SMALL LETTER TRESILLO</li> <li>A72C LATIN CAPITAL LETTER CUATRILLO</li> <li>A72D LATIN SMALL LETTER CUATRILLO</li> <li>A72E LATIN CAPITAL LETTER CUATRILLO WITH COMMA</li> <li>A72F LATIN SMALL LETTER CUATRILLO WITH COMMA</li> </ul> with their glyphs that are similar to those shown in document N3082. </li> </ul>	
aa.	M48.3 (Changes to Vai script): With reference to document N3081 and Irish PDAM3 ballot comment T.3 in document N3071, WG2 accepts to replace the existing Vai block with the charts and names lists as shown in document N3059, which include the	

Item	Assigned to / action (Reference resolutions in document N3104, and unconfirmed minutes in document N3103 for meeting 48 - this document you are reading)	Status
	following changes: c. extending the Vai block by 1 column to end at A62F d. addition of the following fourteen (14) characters to the Vai block: Ten Vai Digits - Zero through Nine and Four new Vai syllables EEN, IN, OON and UN.	
<b>bb.</b>	<i>M48.4 (Change name of one Saurashtra character):</i> With reference to document N3058 and Irish PDAM3 ballot comment T.4 in document N3071, WG2 accepts to change the name of A8B4 in the Saurashtra block: from SAURASHTRA LETTER UPAKSHARA to SAURASHTRA CONSONANT SIGN HAARU.	
<b>cc.</b>	<i>M48.5 (PDAM3 Disposition of Comments):</i> WG2 accepts the disposition of comments on PDAM3 ballot in document N3098 and instructs its editor to prepare the final text of Amendment 3 incorporating the dispositions. The following are noted in particular: e. Changes noted in resolution on 'changes to Mayanist characters' f. Changes noted in resolution on 'changes to Vai script' g. Correct the shape of new A555 VAI SYLLABLE KPAN h. Change noted in resolution on 'change name of one Saurashtra character'	
<b>dd.</b>	<i>M48.6 (Lycian script):</i> WG2 resolves to encode the Lycian script proposed in document N3019 in a new block 10280 -1029F in the SMP, named 'Lycian', and populate it with twenty nine (29) characters with the names, glyphs and code position assignments as shown on pages 10 and 11 in document N3019.	
<b>ee.</b>	<i>M48.7 (Lydian script):</i> WG2 resolves to encode the Right-to-Left Lydian script proposed in document N3019 in a new block 10920 -1093F in the SMP, named 'Lydian', and populate it with twenty seven (27) characters with the names, glyphs and code position assignments as shown on pages 12 and 13 in document N3019.	
<b>ff.</b>	<i>M48.8 (Carian script):</i> WG2 resolves to encode the Carian script proposed in document N3020 in a new block 102A0 -102DF in the SMP, named 'Carian', and populate it with forty nine (49) characters with the names, glyphs and code position assignments as shown on pages 12 and 13 in document N3020.	
<b>gg.</b>	<i>M48.9 (Sundanese script):</i> WG2 resolves to encode the Sundanese script proposed in document N3022 in a new block 1B80 -1BBF in the BMP, named 'Sundanese', and populate it with fifty five (55) characters (some of which are combining) with the names, glyphs and code position assignments as shown on pages 7 and 8 in document N3022.	
<b>hh.</b>	<i>M48.10 (Rejang script):</i> WG2 resolves to encode the Rejang script proposed in document N3096 in a new block A930 -A95F in the BMP, named 'Rejang', in a future amendment to the standard, and populate it with thirty seven (37) characters (some of which are combining) with the names (without the annotations), glyphs and code position assignments as shown on pages 5 and 6 in document N3096.	
<b>ii.</b>	<i>M48.11 (Kayah Li script):</i> WG2 resolves to encode the Kayah Li script proposed in document N3038 in a new block A900 -A92F in the BMP, named 'Kayah Li', and populate it with forty eight (48) characters (some of which are combining) with the names, glyphs and code position assignments as shown on pages 7 and 8 in document N3038.	
<b>jj.</b>	<i>M48.12 (CJK Strokes):</i> With reference to document N3063, WG2 accepts to encode twenty (20) additional characters in code positions 31D0 to 31E3 in the CJK Strokes block, with names, glyphs and code position assignments as shown on page 2 of document N3063.	
<b>kk.</b>	<i>M48.13 (Phaistos disc symbols):</i> WG2 resolves to encode the Phaistos disc symbols proposed in document N3066 in a new block 101D0 - 101FF in the SMP, named 'Phaistos Disc', and populate it with forty six (46) characters (one of which is combining) with the names, glyphs and code position assignments as shown on pages 14 and 15 (excluding two characters at 101FE and 101FF) in document N3066.	
<b>ll.</b>	<i>M48.14 (108 Medievalist characters):</i> With reference to document N3027, WG2 accepts to encode the following in the standard: d. Twenty six (26) characters in the Combining Diacritical Marks Supplement block from pages 52 and 53 (excluding 1DE4) e. Nine (9) characters in the Latin Extended Additional block from pages 54 and 55; with the name change for 1E9F to LATIN SMALL LETTER DELTA f. Seventy three (73) characters in the Latin Extended-D block from pages 58 and 59 (excluding A761 to A766) The final glyphs, names and code position assignments are reflected in the charts in document N3059.	



Item	Assigned to / action (Reference resolutions in document N3104, and unconfirmed minutes in document N3103 for meeting 48 - this document you are reading)	Status
mm.	<p><i>M48.15 (7 Myanmar additions):</i> With reference to documents N3043 and N3099, WG2 accepts the following changes in the Myanmar block:</p> <p>Add the following seven (7) characters:</p> <p>102B MYANMAR VOWEL SIGN TALL AA  103A MYANMAR SIGN ASAT  103B MYANMAR CONSONANT SIGN MEDIAL YA  103C MYANMAR CONSONANT SIGN MEDIAL RA  103D MYANMAR CONSONANT SIGN MEDIAL WA  103E MYANMAR CONSONANT SIGN MEDIAL HA  103F MYANMAR LETTER GREAT SA</p> <p>with their glyphs that are similar to those shown in document N3043.</p> <p>Change the glyphs for the following:</p> <p>104E MYANMAR SYMBOL AFOREMENTIONED  1039 MYANMAR SIGN VIRAMA</p> <p>to the glyphs that are similar to those shown in document N3043.</p>	
nn.	<p><i>M48.16 (14 Myanmar characters for Mon and S'gaw Karen):</i> WG2 accepts to encode in the Myanmar block the fourteen (14) characters, some of which are combining, with their glyphs that are similar to those shown in document N3044:</p> <p>1028 MYANMAR LETTER MON E  1033 MYANMAR VOWEL SIGN MON II  1034 MYANMAR VOWEL SIGN MON O  105A MYANMAR LETTER MON NGA  105B MYANMAR LETTER MON JHA  105C MYANMAR LETTER MON BBA  105D MYANMAR LETTER MON BBE  105E MYANMAR CONSONANT SIGN MON MEDIAL NA  105F MYANMAR CONSONANT SIGN MON MEDIAL MA  1060 MYANMAR CONSONANT SIGN MON MEDIAL LA  1061 MYANMAR LETTER SGAW KAREN SHA  1062 MYANMAR LETTER SGAW KAREN EU  1063 MYANMAR TONE MARK SGAW KAREN HATHI  1064 MYANMAR TONE MARK SGAW KAREN KE PHO</p>	
oo.	<p><i>M48.17 (2 combining marks for Lithuanian phonetics):</i> With reference to document N3048, WG2 accepts the following two (2) combining marks in the Combining Diacritics Supplement block:</p> <p>1DCB COMBINING BREVE-MACRON  1DCC COMBINING MACRON-BREVE</p> <p>with their glyphs that are similar to those shown in document N3048; both are combining characters.</p>	
pp.	<p><i>M48.18 (4 Malayalam characters):</i> With reference to document N3015, WG2 accepts the following four (4) additional Malayalam characters, with their glyphs that are similar to those shown in document N3015 to encode in the Malayalam block:</p> <p>0D44 MALAYALAM VOWEL SIGN VOCALIC RR  0D62 MALAYALAM VOWEL SIGN VOCALIC L  0D3D MALAYALAM PRASLESHAM (avagraha) ←100000th character  0D79 MALAYALAM DATE MARK</p>	
qq.	<p><i>M48.19 (3 Uralic Phonetic characters):</i> With reference to document N3070, WG2 accepts the following three (3) characters in the Latin Extended-C block:</p> <p>2C7B LATIN LETTER SMALL CAPITAL TURNED E  2C7C LATIN SUBSCRIPT SMALL LETTER J  2C7D MODIFIER LETTER CAPITAL V</p> <p>with their glyphs that are similar to those shown in document N3070.</p>	
rr.	<p><i>M48.20 (4 Tibetan characters):</i> WG2 accepts the following four Tibetan characters to encode in the Tibetan block:</p> <p>d. 0FCE TIBETAN SIGN RDEL NAG RDEL DKAR  with its glyph similar to that shown in document N3011</p> <p>e. 0FD2 TIBETAN MARK NYIS TSHEG  with its glyph similar to that shown in document N3032, and,</p> <p>f. 0FD3 TIBETAN MARK INITIAL BRDA RNYING YIG MGO MDUN MA  0FD4 TIBETAN MARK CLOSING BRDA RNYING YIG MGO SGAB MA</p> <p>with their glyphs that are similar to those shown in document N3033,</p>	
ss.	<p><i>M48.21 (3 Phonetic characters for Swedish dialects in Finland):</i> With reference to</p>	



Item	Assigned to / action (Reference resolutions in document N3104, and unconfirmed minutes in document N3103 for meeting 48 - this document you are reading)	Status
	document N3031, WG2 accepts the following three (3) characters to the Latin Extended-C block: 2C78 LATIN SMALL LETTER E WITH TAIL 2C79 LATIN SMALL LETTER TURNED R WITH TAIL 2C7A LATIN SMALL LETTER O WITH LOW RING INSIDE (changed name) with their glyphs that are similar to that shown in document N3031.	
tt.	<i>M48.23 (1 Mongolian character):</i> With reference to document N3041, WG2 accepts the following character in the Mongolian block: 18AA MONGOLIAN LETTER MANCHU ALI GALI LHA with its glyph similar to that shown in document N3041.	
uu.	<i>M48.24 (4 Latin Egyptological characters):</i> With reference to document N2241 and item 4 in document N3045, WG2 accepts the following four (4) characters to the Latin Extended-D block: A722 LATIN CAPITAL LETTER EGYPTOLOGICAL ALEF A723 LATIN SMALL LETTER EGYPTOLOGICAL ALEF A724 LATIN CAPITAL LETTER EGYPTOLOGICAL AIN A725 LATIN SMALL LETTER EGYPTOLOGICAL AIN with their glyphs that are similar to those shown in document N2241.	
vv.	<i>M48.25 (2 Gurmukhi characters):</i> WG2 accepts the following two (2) Gurmukhi characters for encoding in the Gurmukhi block: c. 0A51 GURMUKHI SIGN UDAAT with its glyph similar to that shown in document N3021 d. 0A75 GURMUKHI SIGN YAKASH with its glyph similar to that shown in document N3073; this is a combining character,	
ww.	<i>M48.26 (Named sequences for Lithuanian):</i> <i>Unanimous</i> With reference to document N3090, WG2 accepts the named sequences for composite characters for Lithuanian repertoire identified in the table in Addendum II less the following: Entry 14, which is already encoded as LATIN SMALL LETTER I WITH DOT ABOVE AND ACUTE, 0069 0307 0301 (identical to the requested sequence) in clause 29 Amendment 2 to the standard. WG2 further instructs its project editor to add text to the standard to clarify the rules for adding and naming Named Sequences, with particular attention to those resulting in canonically equivalent composed characters.	
xx.	<i>M48.27 (Extended Collections):</i> WG2 accepts the definition of 'Extended Collection' given in document N3105 for inclusion in the standard.	
yy.	<i>M48.28 (Lithuanian Extended Collection):</i> WG2 accepts the inclusion in the standard of an extended collection named 'Contemporary Lithuanian Letters', an extended collection of standalone characters and named sequences, corresponding to the repertoire shown in Addendum I in document N3090.	
zz.	<i>M48.29 (5 collections corresponding to Japanese standards):</i> With reference to the five requested collections in document N3091, WG2 accepts to define five new collections in the standard based on information provided in attachment (N3091-A) to document N3091.	
aaa.	<i>M48.30 (Hangul clarification):</i> WG2 accepts removing the word 'character' (that occurs three times in the parenthetical notations in the first paragraph) in clause 26.1 in the standard.	
bbb.	<i>M48.32 (Ideographic Variation Database):</i> WG2 accepts the proposed replacement of two paragraphs in clause 20.4 with three new paragraphs and one normative reference to UTS #37, as detailed in document N3035, in order to include by reference the Variation Sequences defined in the Ideographic Variation Database in that standard.	
ccc.	<i>M48.33 (Information on control characters):</i> With reference to document N3046 and to item 4 regarding control characters in ITU-T SG17 request document N3013, WG2 resolves to add to the standard: c. normative definitions for relevant terms related to control characters and control functions d. an informative annex showing the default content of 0000 to 001F, 007F, and 0080 to 009F blocks and charts similar to those published in the Unicode Standard, and names in the form of annotations - "(control) (long name from ISO/IEC 6429)". WG2 further instructs its project editor to prepare appropriate text including the	

Item	Assigned to / action (Reference resolutions in document N3104, and unconfirmed minutes in document N3103 for meeting 48 - this document you are reading)	Status
	feedback from the discussion on document N3059 at meeting 48.	
ddd.	M48.34 (Additions of UTF-16BE and other forms): WG2 accepts the request from ITU-T SG17 in document N3013, item b, to add explicit text for UTF-16BE and other similar forms, and instructs its editor to prepare appropriate text for inclusion in the standard.	
eee.	M48.35 (Updating collection 307 definition): WG2 accepts updating definition of collection '307 Unicode 5.0' to include the four Devanagari characters 097B, 097C, 097E and 097F (encoded in Amendment 3) to correspond to the repertoire of Unicode 5.0.	
fff.	M48.36 (Annotations for 4 Lao characters): With reference to item 6 in document N3045 WG2 accepts to add annotations for the following four (4) characters in the Lao block: 0E9D LAO LETTER FO TAM (fo fon) 0E9F LAO LETTER FO SUNG (fo fay) 0EA3 LAO LETTER LO LING (ro) 0EA5 LAO LETTER LO LOOT (lo).	
ggg.	M48.37 (Progression of Amendment 3): WG2 resolves to include all the items accepted for inclusion in the standard noted in the various resolutions above along with the changes arising out of the disposition of comments into Amendment 3. WG2 instructs its project editor to prepare the revised text of Amendment 3 and forward it along with the disposition of comments document N3098 to the SC2 secretariat for a second PDAM ballot. The final set of charts and names lists are in document N3059. The schedule for this work item is revised accordingly to: PDAM 2006-05-31, FPDAM 2006-10-31, and FDAM 2007-05 (starting dates are given).	
<b>AI-48-4</b>	<b>Ad hoc group on principles and procedures (lead - Dr. V.S. UMAmaheswaran)</b>	
	<i>To take note of and act upon the following items.</i>	
f.	To incorporate into the P&P document appropriate text arising from resolution M48.33 (Information on control characters): With reference to document N3046 and to item a regarding control characters in ITU-T SG17 request document N3013, WG2 resolves to add to the standard: e. normative definitions for relevant terms related to control characters and control functions f. an informative annex showing the default content of 0000 to 001F, 007F, and 0080 to 009F blocks and charts similar to those published in the Unicode Standard, and names in the form of annotations - "(control) (long name from ISO/IEC 6429)". WG2 further instructs its project editor to prepare appropriate text including the feedback from the discussion on document N3059 at meeting 48.	
<b>AI-48-5</b>	<b>IRG Rapporteur (Dr. LU Qin)</b>	
	<i>To take note of and act upon the following items.</i>	
d.	M48.39 (CJK Ext C1): WG2 takes note of item 3 in document N3064, in particular that CJK Ext C1 repertoire is expected to be finalized at IRG meeting 26 in June 2006, and accepts the delay in its submission for WG2 consideration at its next meeting.	
e.	M48.40 (CJK comments on FDAM1): WG2 instructs the IRG to review and prepare a response to the various issues raised regarding CJK Unified Ideographs in the FDAM1 ballot response in document SC2 N3814.	
<b>AI-48-6</b>	<b>TCA (Mr. Shih-Shyeng Tseng)</b>	
b.	To prepare an updated proposal on Mahjong symbols addressing the concerns and feedback received at meeting M48 on document N2975.	
<b>AI-48-7</b>	<b>US national body (Asmus Freytag)</b>	
b.	To prepare updated Arabic Math proposal(s) based on documents N3085 to N3089.	
<b>AI-48-8</b>	<b>National body of Korea (Republic of) (Professor KIM Kyongsok)</b>	
c.	M48.31 (Model for Hangul Jamos): Republic of Korea is invited to prepare a document on 'a new model for using Hangul Jamos' and send to the convener no later than 2006-07-15, for distribution to and review by WG2 experts.	
<b>AI-48-9</b>	<b>All national bodies and liaison organizations</b>	
	<i>To take note of and act upon the following items.</i>	
f.	M48.42 (Future meetings): WG 2 meetings: Meeting 49 – 2006-09-25/29, Tokyo, Japan Meeting 50 – 2007-04-16/20, Frankfurt, Germany Meeting 51 - Fall 2007 - Japan or Hong Kong or Taipei to be confirmed) Meeting 52 - April 2008, Republic of Korea (to be confirmed) IRG meeting:	

Item	Assigned to / action (Reference resolutions in document N3104, and unconfirmed minutes in document N3103 for meeting 48 - this document you are reading)	Status
	IRG #26 - 2006-06-5/9, Hue, Vietnam IRG #27 - 2006-11-27/12-01, Taipei (host: TCA) IRG #28 - 2007 May / June, Hangzhou, China (to be confirmed) IRG#29 - 2007 November, San Jose, California, USA (host Adobe).	
g.	To take note of <i>M48.43 (Roadmap snapshot)</i> : WG2 instructs its convener to post an updated snapshot of the roadmaps (in document N3005) to the WG2 web site.	
h.	To take note of revised Principles and Procedures document N3002 - revised formatting of proposal summary form.	
i.	In addition to the proposals mentioned in the items above, the following proposals have been carried forward from earlier meetings. All national bodies and liaison organizations are invited to review and comment on them. <ul style="list-style-type: none"> <li>r. Manichaeian script – document N2544</li> <li>s. Avestan and Pahlavi – document N2556</li> <li>t. Dictionary Symbols – document N2655</li> <li>u. Invisible Letter – document N2822</li> <li>v. Palestinian Pointing characters - document N2838</li> <li>w. Babylonian Pointing characters – document N2759</li> <li>x. Bantu Phonetic Click characters – document N2790</li> <li>y. Samaritan Pointing - 12 Hebrew characters – document N2758</li> <li>z. Dominoe Tiles and other Game symbols – N2760</li> <li>aa. Four Tibetan charactes for Balti -- N3010</li> <li>bb. Kaithi script -- N3014</li> <li>cc. Karen, Shan and Kayah characters -- N3080</li> <li>dd. Arabic Math proposals -- N3085, N3086, N3087, N3088 and N3089Cyrillic combining marks -- N3097</li> </ul>	

=====  
 ===== End of Document =====