



**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 50  
German National Library, Frankfurt-am-Main, Germany; 2007-04-24/27

**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 2007-08-31.

**ACTION ID:** ACT

**DUE DATE:** 2007-08-31

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

**MEDIUM:** Acrobat PDF file

**NO. OF PAGES:** 53 (including cover sheet)

Michael Y. Ksar  
Convener – ISO/IEC/JTC 1/SC 2/WG 2  
22680 Alcalde Rd  
Cupertino, CA 95014  
U.S.A.

Phone: +1 408 255-1217  
Email: [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 **N3253**  
**2007-07-26**

<b>Title:</b>	Unconfirmed minutes of WG 2 meeting 50 German National Library, Frankfurt-am-Main, Germany; 2007-04-23/27
<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@10646.com">mikeksar@10646.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

Input document:

[3199](#) 1st Call for meeting 51 - Frankfurt; Ksar; 2007-02-08

The meeting was convened at 10:10h.

Mr. Mike Ksar: I welcome you all to Frankfurt. I appreciate your contributions and effort. We have several invited experts here - some are not part of the national body delegations. Our host Ms. Ute Schwens, Director of German National Library will address you now.

Ms. Ute Schwens: Our warm welcome to you to the German National Library. As director of this location I am proud to have the honour of hosting the 50th meeting of UCS working group. Last week we held a meeting of about 250 delegates. This week we have you here. These are good examples of international standardization and cooperation. ISO/IEC 10646 is one of the best examples of such cooperation -- containing 10000+ characters, meeting the user needs of the world's languages. German language has several needs for its contents; not to speak of the different scripts in the German National Libraries; we have to catalog the various scripts. You helped us to distinguish between Trema and Umlaut. This week we will have the discussion on German sharp S. You will have interesting conversations with the experts. We would like to show you the work we do in the library. We wish you a successful meeting and hope you have a nice stay in Frankfurt.

Mr. Paul Dettmer: Some of you know me from former meetings. On behalf of DIN and German National Library I welcome you all to the 50th WG2 meeting in Frankfurt. I wish you a pleasant stay in Germany and wish you a successful meeting. Mr. Reinhold Heuvelmann has organized the meeting facilities. He will explain the logistics.

Mr. Reinhold Heuvelmann: I am with the German National Library. Hopefully this meeting room is large enough for all the guests. I ask for your patience. We have another room for about 24 participants available down the hall. We have wireless LAN in the building. It is working for most of you. The badges that you have are also visitor cards. Please keep them with you for the week. We have made 10 sets of two volumes of printout of the meeting documents. There are 2 printer stations outside this room. You can use memory sticks for printing. There is a restaurant in the ground floor and there are several restaurants in the area. There is coffee and water just outside the meeting room. Smoking is not permitted inside the building. As for the social event, we are organizing a boat trip on Wednesday afternoon or evening. Additional information is provided in your booklet. Please let me know if you want to take part in this event. As Ms. Ute Schwens mentioned, we have two guided tours of the library. On Friday afternoon - a one-hour tour of the German National Library is being organized. Two groups starting at 1:00 pm and at 1:30 pm. Whoever is interested can meet in the lobby downstairs. If you need any assistance please ask me.

Mr. Mike Ksar: In the brochure we have the information about the social event. It will cost the delegates 54 Euros without the drinks. We need to know who will attend before end of today. In attempting to reduce the costs of the meetings the hosting organizations are not required to provide any social events. It is up to the delegates to pay for the event.

A hardcopy of the latest agenda has been made available. There is a CD containing all the documents that are on the agenda. You can copy them to your personal computers. I will circulate it around. If you have any new contributions please give them to me. If we have time to discuss them this at this meeting we will. The deadline for contributions for this meeting is past. I will review each contribution and post it if it is technical and if it is suitable for WG2 discussion. The meeting agenda is broken into several parts. It follows the usual formats we have been following. Hardcopies of documents containing the draft disposition of ballot comments will be made available at the end of the day.

## 1.1 Roll Call

Input document:

[3151](#) WG2 Experts list - Post Meeting 49; Ksar; 2007-03-08

A document containing the names and other contact details of WG2 experts was circulated. Attendees were requested to make any corrections and mark their attendance in that document and to give their business card to the recording secretary. The following 44 delegates representing 11 national bodies, 3 liaison organizations, SC2 secretariat, including 5 invited experts were present at different times during the meeting.

Name	Representing	Affiliation
Mike KSAR	.Convener, USA	Independent
Ute SCHWENS	.Host, Germany	Director, German National Library
Johannes BERGERHAUSEN	.Invited Expert, Germany	University of Applied Sciences, Mainz
Malcolm HYMAN	.Invited Expert, Germany	Max-Planck Institute for the History of Science, Berlin
Bruce PATERSON	.Invited Expert, UK	Previous editor of ISO/IEC 10646
Peter CONSTABLE	.Invited Expert, USA	Microsoft Corporation
Peter M. SCHARF	.Invited Expert, USA	Dept. of Classics, Brown University
Alain LABONTÉ	Canada Editor 14651	Ministère des services gouvernementaux du Québec
V. S. (Uma) UMAMAHESWARAN	Canada Recording Secretary	IBM Canada Ltd.
Azhati PIREDEWUSI	China	Xinjiang University
CHEN Zhuang	China	Chinese Electronics Standardization Institute
DAI Hong	China	Chinese Electronics Standardization Institute
DALUOSANGLANGJIE	China	Tibet University
Lobsang TOKME	China	Tibet Autonomous Region
NGODRUP	China	Tibet University
TAN Yuting	China	Yunnan Ethnic Affairs Committee
WANG Shizhong	China	National Culture Research Institute, Bijie University
WU Xuejun	China	Ancient Yi Text Research Institute, Bijie University
Wushour SILAMU	China	Xinjiang University
ZHANG Liwei	China	Zhonghua Book Company
Anssi YLI-JYRÄ	Finland	CSC - Scientific Computing Ltd.
Andreas STÖTZNER	Germany	Signography
Mark KÜSTER	Germany	University of Applied Sciences, Worms
Paul DETTMER	Germany	DIN
Reinhold HEUVELMANN	Germany	German National Library
Michael EVERSON	Ireland Contributing Editor	Evertime
LU Qin	IRG Rapporteur	Hong Kong Polytechnic University
Masahiro SEKIGUCHI	Japan	Fujitsu Ltd.
Yasuhiro ANAN	Japan	Microsoft Japan
Dae Hyuk AHN	Korea (Republic of)	Microsoft Korea
KIM Kyongsok	Korea (Republic of)	Pusan National University
Zong Su KIM	Korea (Republic of)	Department of Korean Linguistics and Literature, Hanyang University
Teodor STATESCU-NEGULESCU	Romania	Romanian Standards Association
Tatsuo KOBAYASHI	SC2 Chair, Japan	Justsystem Corporation

Name	Representing	Affiliation
Lin-Mei WEI	TCA - Liaison	Chinese Foundation Of Digitization Technology
Shih-Shyeng TSENG	TCA - Liaison	Academia Sinica
Andrew WEST	UK	Independent
Martin HOSKEN	UK	SIL International, Thailand
Ken WHISTLER	USA Contributing Editor	Sybase Inc.
Michel SUIGNARD	USA Project Editor	Microsoft Corporation
Asmus FREYTAG	USA The Unicode Consortium - Liaison Contributing Editor	Unicode Inc.
Deborah ANDERSON	USA; SEI, UC Berkeley - Liaison	Dept. of Linguistics, UC Berkeley
Thiem CAM BACH	Vietnam	Centre for Ethnic Education Studies, Ministry of Education and Training
Viet NGO TRUNG	Vietnam	Institute of Information Technology - VAST

Drafting Committee: The recording secretary Dr. Umamaheswaran was assisted by Ms. Deborah Anderson, Messrs. Ken Whistler, Mike Ksar, Michael Everson, Masahiro Sekiguchi and Michel Suignard in drafting the meeting resolutions.

## 2 Approval of the agenda

Input document:

[3205](#) Proposed agenda for Meeting 50; Ksar; 2007-04-21

Mr. Mike Ksar: Several contributions were received quite late. All the documents are posted on the WG2 web site. We can review the agenda in document [N3205](#) and ensure all the documents are assigned to different agenda items.

### Discussion:

- a. Dr. Ken Whistler: Document [N3237](#) should be moved from 8.17 to 8.5. The two agenda items are duplicates. The agenda items 9.2 to 9.5 are all for Amd. 4, and should be under item 8.
- b. Mr. Mike Ksar: Item 8.17 will be merged with item 8.5. Items 9.2 to 9.5 will become items 8.22 to 8.25 respectively.
- c. Dr. Ken Whistler: Under the new item 8.25; three more documents [N3216](#), [N3217](#) and [N3218](#) should be added for discussion.
- d. Mr. Mike Ksar: There is also a new Orkhon contribution - document [N3258](#) from Professor Woshour Silamu of Xinjiang University.
- e. Mr. Yasuhiro Anan: We need a new item 9.3 for 7 CJK additions in document [N3210](#).

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. Some items that were not discussed have been deleted. Some items have been merged under other items where the associated contributions were discussed. The following table of contents reflects where the items are discussed and recorded.

Item Number	Title	Page
1	Opening	2
1.1	Roll Call	3
2	Approval of the agenda	4
3	Approval of minutes of meeting 49	6
4	Review action items from previous meeting	6
4.1	Outstanding action items from meeting 47, 2005-09-12/15, Sophia Antipolis, France	6
4.2	Outstanding action items from meeting 48, 2006-04-24/27, Mountain View, CA, USA	6
4.3	New action items from meeting 49, AIST, Akihabara, Tokyo, Japan; 2006-09-25/29	6
5	JTC1 and ITTF matters	11
6	SC2 matters	11
6.1	SC2 Program of Work	11
6.1.1	Liaison from JTC1/SC6 - future editions	11
6.1.2	SC2 Proposed Business Plan to JTC1	11
6.2	Ballot results – FPDAM 3 and PDAM 4	11
7	WG2 matters/housekeeping	12

Item Number	Title	Page
7.1	Principles & Procedures	12
7.2	Roadmap	12
7.3	Request from SEI - UC Berkeley re: Minority Scripts	12
8	Scripts contributions – not part of PDAM 4	12
8.1	Avestan script	12
8.2	Bopomofo characters	12
8.2.1	Bopomofo Letter IH	12
8.2.2	Bopomofo Letter YI	13
8.3	8 Arabic characters for Persian and Azerbaijani	14
8.4	4 Qur'anic Arabic letters	14
8.5	Egyptian Hieroglyphs	15
8.6	Additional Cyrillic characters	15
8.7	Medievalist and Iranianist punctuation characters	16
8.8	Archaic Sinhala numbers	17
8.9	29 additional Mathematical and Symbol characters	17
8.10	Meitei Mayek script	18
8.11	Bamum script	19
8.12	Tai Viet script	20
8.13	Additions for Coptic and Latin	20
8.14	Lanna script	21
8.15	Characters for Vedic Sanskrit	23
8.16	Anatolian Hieroglyphs	25
8.17	Parthian, Inscriptional Pahlavi and Psalter Pahlavi scripts	25
8.18	Additional historic syllables for Vai	25
8.19	Devanagari letter Candra A	26
8.20	Latin Capital Letter Sharp S	26
8.21	3 Additional Tibetan characters	28
8.22	Additional Hangul Jamos	28
8.23	Proposed additions to ISO/IEC 10646:2003	29
8.23.1	Additional Orthographic and Modifier Characters	29
8.23.2	Symbol for Samaritan Source	29
8.23.3	Glyph corrections - 027F and 0285	30
8.23.4	Addition of 11 Ancient Roman Characters	30
8.24	Orkhon script	30
9	Other contributions	31
9.1	Proposal to disunify U+4039	31
9.2	Ideographic Variation Database - referencing	33
10	Proposed Disposition of Comments	33
10.1	FPDAM 3 ballot	33
10.2	PDAM 4 ballot	35
11	Publication issues	37
11.1	Draft results of repertoire review for FPDAM 3, FPDAM4 and future amendments	37
11.2	About the Code Table Format	37
11.3	New edition draft	39
12	IRG status and reports	42
12.1	IRG Meeting 27	42
12.2	CJK Ext C	43
12.3	Addition of 7 CJK Unified Ideographs	44
13	Defect reports	45
14	Liaison & national body reports	45
14.1	Unicode Consortium	45
14.2	SEI – UC Berkeley	45
15	Other business	45
15.1	Web Site	45
15.2	Future Meetings	46
16	Closing	46
16.1	Approval of Resolutions of Meeting 50	46
16.2	Adjournment	47
17	Action Items	47
17.1	Outstanding action items from meeting 47, 2005-09-12/15, Sophia Antipolis, France	47
17.2	Outstanding action items from meeting 48, 2006-04-24/27, Mountain View, CA, USA	47
17.3	New action items from meeting 50, 2007-04-23/27, Frankfurt-Am-Main, Germany	47

### 3 Approval of minutes of meeting 49

Input document:

[3153](#) Minutes Meeting 49; Uma/Ksar; 2007-02-16

Dr. Umamaheswaran presented the minutes briefly, mentioning that he has not received any feedback on the posted minutes.

Dr. Asmus Freytag: It is up to the usual high standards of quality and I did not find any issues with it.

**Disposition:** The meeting minutes were adopted as written.

### 4 Review action items from previous meeting

Input document:

[3153AI](#) Sections 16 of minutes; action items out of M49; Uma/Ksar; 2007-02-16

Dr. Umamaheswaran reviewed the outstanding action items in document N[3153](#)-AI. 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 M48, and new action items from the latest meeting M49 are listed in the tables in the following sections.

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

Item	Assigned to / action (Reference resolutions in document N <a href="#">2954</a> , and unconfirmed minutes in document N <a href="#">2953</a> for meeting 47 - with any corrections noted in section 3 of document N <a href="#">3103</a> 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>	
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. M48, M49, M50 - in progress.	In progress.

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

Item	Assigned to / action (Reference resolutions in document N <a href="#">3104</a> , and unconfirmed minutes in document N <a href="#">3103</a> for meeting 48 - with any corrections noted in section 3 of document N <a href="#">3153</a> from meeting 49).	Status
AI-48-4	<b>Ad hoc group on principles and procedures (lead - Dr. V.S. UMaaheswaran)</b>	
	<i>To take note of and act upon the following items.</i>	
b.	To incorporate into the P&P document appropriate text arising from resolution M48.33 ( <i>Information on control characters</i> ): With reference to document N <a href="#">3046</a> and to item a regarding control characters in ITU-T SG17 request document N <a href="#">3013</a> , WG2 resolves to add to the standard: <ol style="list-style-type: none"> <li>a. normative definitions for relevant terms related to control characters and control functions</li> <li>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)".</li> </ol> WG2 further instructs its project editor to prepare appropriate text including the feedback from the discussion on document N <a href="#">3059</a> at meeting 48. M48, M49 - in progress.	Completed; no effect on P&P document
AI-48-7	<b>US national body (Asmus Freytag)</b>	
b.	To prepare updated Arabic Math proposal(s) based on documents N <a href="#">3085</a> to N <a href="#">3089</a> . M48, M49, M50 - in progress.	In progress

#### 4.3 New action items from meeting 49, AIST, Akihabara, Tokyo, Japan; 2006-09-25/29

Item	Assigned to / action (Reference resolutions in document N <a href="#">3154</a> , and unconfirmed minutes in document N <a href="#">3153</a> for meeting 49.)	Status
AI-49-1	<b>Meeting Secretary - Dr. V.S. UMaaheswaran</b>	
a.	To finalize the document N <a href="#">3154</a> containing the adopted meeting resolutions and send it to the convener as soon as possible.	Completed. See document N <a href="#">3154</a> .

Item	Assigned to / action (Reference resolutions in document <a href="#">N3154</a> , and unconfirmed minutes in document <a href="#">N3153</a> for meeting 49.)	Status
b.	To finalize the document <a href="#">N3153</a> containing the unconfirmed meeting minutes and send it to the convener as soon as possible.	Completed. See document <a href="#">N3153</a> .
<b>AI-49-2</b>	<b>Convener - Mr. Mike Ksar</b>	
a.	<i>M49.30 (Roadmap snapshot)</i> : WG2 instructs its convener to post the updated snapshot of the roadmaps (in document <a href="#">N3149</a> ) to the WG2 web site.	Completed. Document <a href="#">N3149</a> is posted.
b.	To add the following carried forward scripts to next meeting agenda, if any updates are available: Manichaeen script – document <a href="#">N2544</a> Avestan and Pahlavi script– document <a href="#">N2556</a> Dictionary Symbols – document <a href="#">N2655</a> Samaritan Pointing characters – document <a href="#">N2758</a> Babylonian Pointing characters – document <a href="#">N2759</a> Bantu Phonetic Click characters – document <a href="#">N2790</a> Invisible Letter – document <a href="#">N2822</a> Palestinian Pointing characters - document <a href="#">N2838</a> Kaithi script -- document <a href="#">N3014</a> Arabic Math Symbols -- documents <a href="#">N3085</a> , <a href="#">N3087</a> , <a href="#">N3088</a> and <a href="#">N3089</a> Anatolian Hieroglyphs-- document <a href="#">N3144</a> Meithei Mayek script -- document <a href="#">N3158</a> Orkhon script - document <a href="#">N3164</a> Hangul syllables -- documents <a href="#">N3168</a> and <a href="#">N3172</a> .	Avestan, Anatolian Hieroglyphs, Egyptian Hieroglyphs, Pahlavi, Meithei Mayek, Orkhon and Hangul are on M50 agenda.
<b>AI-49-3</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 <a href="#">N3186</a> and <a href="#">N3208</a> .

Item	Assigned to / action (Reference resolutions in document <a href="#">N3154</a> , and unconfirmed minutes in document <a href="#">N3153</a> for meeting 49.)	Status
a.	<p>M49.1 (disposition of PDAM3.2 ballot comments): WG2 accepts the disposition of ballot comments on PDAM3.2 in document <a href="#">N3165</a> and instructs its editor to prepare the final text of Amendment 3 incorporating the dispositions. The following changes are noted in particular:</p> <ul style="list-style-type: none"> <li>a. Add 1035 MYANMAR VOWEL SIGN E ABOVE in Myanmar block, with glyph shown in document <a href="#">N3115</a>.</li> <li>b. Change the name for (the 10000th character in ISO/IEC 10646): 0D3D MALAYALAM PRASLESHAM (avagraha) to 0D3D MALAYALAM SIGN AVAGRAHA (praslesham).</li> <li>c. Change the glyph for 0373 GREEK SMALL LETTER ARCHAIC SAMPI to the one shown on the right under Irish comment E.1 in document <a href="#">N3145</a>.</li> <li>d. Change the glyphs for the range of code positions A722 to A725 as shown in Irish comment E.3 in document <a href="#">N3145</a>.</li> <li>e. Add annotations "da nying yik go dun ma" and "da nying yik go kab ma" to the names for two Tibetan characters 0FD3 and 0FD4 respectively.</li> <li>f. Correct the names for the named sequences - changing CAPITAL to SMALL - for:  &lt;012F, 0307, 0301&gt; LATIN SMALL LETTER I WITH OGONEK AND DOT ABOVE AND ACUTE  &lt;012F, 0307, 0303&gt; LATIN SMALL LETTER I WITH OGONEK AND DOT ABOVE AND TILDE.</li> <li>g. Correct the glyphs for 0485 COMBINING CYRILLIC DASIA PNEUMATA and 0486 COMBINING CYRILLIC PSILI PNEUMATA based on document <a href="#">N3118</a>.</li> <li>h. Correct the glyphs for 0340 COMBINING GRAVE TONE MARK and 0341 COMBINING ACUTE TONE MARK to be the same as their canonical equivalent characters at 0300 and 0301 respectively.</li> </ul>	
b.	<p>M49.2 (Telugu additions): With reference to documents <a href="#">N3126</a>, <a href="#">N3156</a>" and item 10 in document <a href="#">N3129</a>, WG2 accepts to encode in the standard 14 additional characters in the Telugu block, at code positions - 0C3D, 0C58, 0C59, 0C62, 0C63, 0C71, and the range 0C78 to 0C7F, with their names and glyphs as shown in document <a href="#">N3116</a>.</p>	
c.	<p>M49.3 (Greek and Latin additions): With reference to document <a href="#">N3122</a> and item 11 in document <a href="#">N3129</a>, WG2 accepts to encode in the standard 1 Greek and 16 Latin characters, with their names, code positions and glyphs as shown in document <a href="#">N3122</a>.</p>	
d.	<p>M49.4 (Tibetan characters for Balti): With reference to document <a href="#">N2985</a>, WG2 accepts to encode in the standard 2 additional characters in the Tibetan block:  0F6B TIBETAN LETTER KKA, and  0F6C TIBETAN LETTER RRA  with their glyphs as shown in document <a href="#">N2985</a>.</p>	
e.	<p>M49.5 (Miscellaneous): WG2 accepts to encode in the standard the following 3 additional characters:</p> <ul style="list-style-type: none"> <li>a. A788 MODIFIER LETTER LOW CIRCUMFLEX ACCENT with its glyph as shown in document <a href="#">N3140</a>.</li> <li>b. 0971 DEVANAGARI SIGN HIGH SPACING DOT with its glyph as shown in document <a href="#">N3125</a>.</li> <li>c. 0BD0 TAMIL OM with its glyph as shown in document <a href="#">N3119</a>.</li> </ul>	
f.	<p>M49.6 (Astrological symbols): With reference to document <a href="#">N3110</a> and item 4 in document <a href="#">N3129</a>, WG2 accepts to encode in the standard 10 Astrological symbols at code positions 26B3 to 26BC in the Miscellaneous Symbols block with their names and glyphs as shown in document <a href="#">N3110</a>.</p>	
g.	<p>M49.7 (Arabic Math symbols): With reference to document <a href="#">N3086</a> and item 3 in document <a href="#">N3129</a>, WG2 accepts to encode:</p> <ul style="list-style-type: none"> <li>a. 21 characters at code positions 2B30 to 2B44 in the Miscellaneous Symbols and Arrows block,</li> <li>b. 5 characters at code positions 0606 to 060A in the Arabic block, and</li> <li>c. 269D OUTLINED WHITE STAR in the Miscellaneous Symbols block,</li> </ul> with their names and glyphs as shown in document <a href="#">N3086</a> .	



Item	Assigned to / action (Reference resolutions in document <a href="#">N3154</a> , and unconfirmed minutes in document <a href="#">N3153</a> for meeting 49.)	Status
h.	<i>M49.8 (Arabic characters for Khojar, Torwali and Burushaski):</i> With reference to document <a href="#">N3117</a> and item 2 in document <a href="#">N3129</a> , WG2 accepts to encode 16 additional characters at code positions 076E to 077D in the Arabic Supplement block with their names as shown in document <a href="#">N3129</a> and their glyphs as shown in document <a href="#">N3117</a> .	
i.	<i>M49.9 (Additional Updates to Amd. 3):</i> WG2 accepts to include in the standard the changes proposed by the project editor in document <a href="#">N3148</a> for the following items: <ul style="list-style-type: none"> <li>a. Usage of the term 'composite characters',</li> <li>b. Referencing latest version of Unicode in clause 20.4,</li> <li>c. Rewording of Note 1 under clause 25,</li> <li>d. Renumbering clauses regarding Name Uniqueness, and</li> <li>e. Synchronizing the clause on Mirrored Characters with the Unicode Standard.</li> </ul>	
j.	<i>M49.10 (Corrections to add to Amd. 3):</i> With reference to the two errors related to new JIS collections reported in document <a href="#">N3166</a> , WG2 accepts for inclusion in the standard the corrections proposed in that document.	
k.	<i>M49.11 (Progression of Amendment 3):</i> WG2 resolves to include all the items accepted for inclusion in the standard noted in resolutions M49.1 to M49.10 above, along with the changes arising out of the disposition of comments, into Amendment 3. WG2 instructs its project editor to forward the final text of Amendment 3 along with the disposition of comments document <a href="#">N3165</a> to the SC2 secretariat for an FPDAM ballot. The final set of charts and names lists are in document <a href="#">N3173</a> . The revised starting dates for this work item are: FPDAM 2006-11-30, and FDAM 2007-06.	
l.	<i>M49.12 (Updates for Amendment 4):</i> WG2 accepts to amend the standard based on the changes proposed by the project editor in document <a href="#">N3148</a> for the following items: <ul style="list-style-type: none"> <li>a. Fixing errors in CJK Source References based on feedback from IRG in document <a href="#">N3132</a>, and</li> <li>b. Text towards deprecating Levels 1 and 2 in the standard.</li> </ul>	
m.	<i>M49.13 (Old Cyrillic):</i> With reference to document <a href="#">N3097</a> and item 1 in document <a href="#">N3129</a> , WG2 accepts to encode in the standard 22 combining characters, at code positions 2DE0 to 2DF5 in a new block 2DE0 -- 2DFF named Cyrillic Extended-A, with their names from document <a href="#">N3129</a> and glyphs from document <a href="#">N3097</a> .	
n.	<i>M49.14 (Game symbols):</i> With reference to document <a href="#">N3171</a> , WG2 accepts to encode in the standard the following: <ul style="list-style-type: none"> <li>a. 44 Mahjong Tiles symbols in a new block 1F000 -- 1F02F named Mahjong Tiles, with the code positions, character names and glyphs as proposed in document <a href="#">N3171</a>,</li> <li>b. 100 Domino Tiles symbols in a new block 1F030 -- 1F09F named Domino Tiles, and</li> <li>c. 4 Draughts symbols in the Miscellaneous Symbols block</li> </ul> with their code positions, character names and glyphs as proposed in document <a href="#">N3171</a> .	
o.	<i>M49.15 (Myanmar additions for Shan and Palaung):</i> With reference to document <a href="#">N3143</a> , WG2 accepts to encode in the standard 23 additional Myanmar characters, some of which are combining characters, at code positions 1022, and 1075 to 108A, with their glyphs and character names as proposed in document <a href="#">N3143</a> .	
p.	<i>M49.16 (Myanmar additions for Karen and Kayah):</i> With reference to document <a href="#">N3142</a> , WG2 accepts to encode in the standard 16 additional Myanmar characters, some of which are combining characters, for Karen and Kayah languages, at code positions 1065 to 1074, with their glyphs and character names as proposed in document <a href="#">N3142</a> .	
q.	<i>M49.17 (Lanna script):</i> With reference to document <a href="#">N3121</a> and the ad hoc report in document <a href="#">N3169</a> on the Lanna script, WG2 accepts to encode in the standard the 127 characters, several of which are combining characters, with their names, glyphs, and code positions as proposed in document <a href="#">N3121</a> in a new block 1A20 -- 1AAF named Lanna.	
r.	<i>M49.18 (Cham script):</i> With reference to document <a href="#">N3120</a> on Cham script, WG2 accepts to encode in the standard the 83 characters, several of which are combining characters, with their names, glyphs, and code positions as proposed in document <a href="#">N3120</a> in a new block AA00 -- AA5F named Cham.	

Item	Assigned to / action (Reference resolutions in document <a href="#">N3154</a> , and unconfirmed minutes in document <a href="#">N3153</a> for meeting 49.)	Status
s.	<i>M49.19 (CJK Unified Ideographs Extension C)</i> : With reference to document <a href="#">N3134</a> , WG2 accepts to encode in the standard the 4219 CJK unified ideographs in a new block 2A700 -- 2B77F named CJK Unified Ideographs Extension C, at code positions 2A700 to 2B77A, in a multiple-column format, similar to the current format for CJK Unified Ideographs, with the glyphs from document <a href="#">N3134</a> .	
t.	<i>M49.20 (Ancient Roman symbols)</i> : With reference to document <a href="#">N3138</a> and item 9 in document <a href="#">N3129</a> , WG2 accepts to encode in the standard 11 characters at code positions 10190 to 1019A in a new block 10190 -- 101CF named Ancient Symbols, with their names and glyphs as shown in document <a href="#">N3138</a> .	
u.	<i>M49.21 (Malayalam Chillus)</i> : With reference to document <a href="#">N3126</a> and item 12 in document <a href="#">N3129</a> , WG2 accepts to encode in the standard the 6 additional characters to the Malayalam block, with their code positions and names as shown in document <a href="#">N3129</a> and glyphs as shown in document <a href="#">N3126</a> .	
v.	<i>M49.22 (Amendment 4 – subdivision and PDAM text)</i> : WG2 instructs its editor to prepare a project sub division proposal and PDAM text based on resolutions M49.12 to M49.21 above, and forward them to the SC2 secretariat for ballot. The proposed completion dates for the progression of this work item are: PDAM 2006-12-15, FPDAM 2007-06-15, and FDAM 2007-12.	
<b>AI-49-4</b>	<b>Ad hoc group on principles and procedures (lead - Dr. V.S. UMaaheswaran)</b>	
	<i>To take note of and act upon the following items.</i>	
a.	<i>M49.26 (Principles and Procedures)</i> : With reference to document <a href="#">N3130</a> , WG2 accepts removal of the question on Levels of Implementation from the Proposal Summary Form. In addition WG2 instructs its ad hoc on Principles and Procedures to formulate appropriate text regarding "what is appropriate for inclusion in ballot comments" accommodating the comments received from this meeting. Further WG2 instructs its ad hoc on P&P to post the updated Principles and Procedures document to the WG2 site.	Completed; see document <a href="#">N3102</a> .
<b>AI-49-4</b>	<b>Irish national body - Mr. Michael Everson</b>	
a.	Mr. Michael Everson is to send feedback to the authors of document <a href="#">N3116</a> requesting more information on the proposed Telugu characters <i>Abbreviation Sign</i> and <i>Talakattu</i> .	Dropped.
<b>AI-49-5</b>	<b>All national bodies and liaison organizations</b>	
	<i>To take note of and act upon the following items.</i>	
a.	<i>M49.29 (Future meetings)</i> : WG 2 meetings: Meeting 50 - 2007-04-23/27, Frankfurt, Germany (backup: USA) Meeting 51 - 2007-09-17/21, Urumqi, Xinjiang, China (backup: Seoul, Republic of Korea) along with SC2 plenary. Meeting 52 - Spring 2008, Taipei (TCA) (backup: Seoul, Republic of Korea) Meeting 53 - Fall 2008 - London, UK (pending confirmation) along with SC2 plenary. IRG meeting: IRG #27 - 2006-11-27/12-01, Sanya, Hainan (China) - (changed location) IRG #28 - 2007-06-4/8, Taipei, Taiwan (TCA) IRG #29 - 2007-11-12/16, San Jose, California, USA (host Adobe; to be confirmed). IRG #30 - June 2008, Busan (Republic of Korea) (tentative) IRG #31 - November 2008, Kunming, Yunnan (China) (tentative) IRG #32 - June 2009 (seeking host).	Noted.
b.	<i>M49.23 (Hangul proposals)</i> : With reference to documents <a href="#">N3168</a> and <a href="#">N3172</a> on old Hangul Jamos, WG2 encourages the ad hoc on Hangul Syllables to continue their discussion regarding suitable text for clause 26.1. National bodies and liaison organizations are requested to review and provide feedback on documents <a href="#">N3168</a> and <a href="#">N3172</a> , for consideration at meeting M50.	Completed; see documents <a href="#">N3242</a> and <a href="#">N3257</a> .
c.	<i>M49.24 (Orkhon proposal)</i> : With reference to document <a href="#">N3164</a> WG2 requests national bodies and liaison organizations to review and provide feedback on the contribution for consideration at meeting M50.	Completed. see document <a href="#">N3258</a> .

Item	Assigned to / action (Reference resolutions in document <a href="#">N3154</a> , and unconfirmed minutes in document <a href="#">N3153</a> for meeting 49.)	Status
d.	<p>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.</p> <p>Manichaean script – document <a href="#">N2544</a>  Avestan and Pahlavi script– document <a href="#">N2556</a>  Dictionary Symbols – document <a href="#">N2655</a>  Samaritan Pointing characters – document <a href="#">N2758</a>  Babylonian Pointing characters – document <a href="#">N2759</a>  Bantu Phonetic Click characters – document <a href="#">N2790</a>  Invisible Letter – document <a href="#">N2822</a>  Palestinian Pointing characters - document <a href="#">N2838</a>  Kaithi script -- document <a href="#">N3014</a>  Arabic Math Symbols -- documents <a href="#">N3085</a>, <a href="#">N3087</a>, <a href="#">N3088</a> and <a href="#">N3089</a>  Anatolian Hieroglyphs-- document <a href="#">N3144</a>  Meithei Mayek script -- document <a href="#">N3158</a></p>	Avestan, Anatolian Hieroglyphs, Egyptian Hieroglyphs, Pahlavi, and Meithei Mayek are on M50 agenda.

## 5 JTC1 and ITTF matters

None for this meeting.

## 6 SC2 matters

### 6.1 SC2 Program of Work

#### 6.1.1 Liaison from JTC1/SC6 - future editions

Input document:

[3112](#) Liaison Contribution to ISO/IEC JTC 1/SC 2 on requirements for future editions of ISO/IEC 10646; JTC1/SC6; 2006-06-22

Mr. Tatsuo Kobayashi: I met with the SC6 chair and ITU/SG representative. The liaison statements were written by the same persons on their side. We had invited them to participate. They have not responded.

#### Discussion:

- a. Dr. Umamaheswaran: We discussed the same document at the last WG2 meeting.
- b. Mr. Mike Ksar: We have considered this document in Tokyo. WG2 has sent the feedback.
- c. Mr. Tatsuo Kobayashi: I will ask Ms. Toshiko Kimura to remind them again.

#### 6.1.2 SC2 Proposed Business Plan to JTC1

Mr. Mike Ksar: Hopefully the resolutions from this meeting will assist our chair to prepare some input to JTC1.

Mr. Tatsuo Kobayashi: The JTC1 plenary is in September 2007. The SC2 plenary is in October. If WG2 and OWG-SORT have some items to carry forward to JTC1, include them in your resolutions.

## 6.2 Ballot results – FPDAM 3 and PDAM 4

Input documents:

[3186](#) FPDAM3 ballot text; FPDAM3 names; Ideographic Extensions; Michel Suignard; 2006-11-30

[3208](#) PDAM4; Project Editor; 2006-12-21

[3223](#) FPDAM 3 Summary of Voting/Table of Replies; SC2 N3924; SC2 Secretariat; 2007-04-05

[3224](#) PDAM 4 Summary of Voting/Table of Replies; SC2 N3920; SC2 Secretariat; 2007-03-15

Mr. Mike Ksar: The ballot responses are in document [N3223](#) and [N3224](#). They will be discussed when we go through the draft disposition of comments in documents [N3225](#) and [N3226](#). (20 hard copies of the draft disposition documents were requested to be made available for the delegates.)

From document [N3223](#) on FPDAM3, the results were: of the 31 P members of SC2, 14 had not voted, 5 had abstained, 2 (Ireland and Japan) had disapproved and 10 had approved (UK and USA had comments).

From document [N3224](#) on PDAM4, the results were: of the 31 P members of SC2, 15 had not voted, 3 had abstained, 4 (China, Japan, UK and USA) had disapproved, and 9 had approved (Ireland had comments).

## 7 WG2 matters/housekeeping

### 7.1 Principles & Procedures

Input document:

[3102](#) Updated P&P; Uma - Ad hoc group on Principles & Procedures; 2007-03-14

Dr. Umamaheswaran: Document [N3102](#) has been posted to the WG2 site. It incorporates the outcome from meeting 49 discussions.

### 7.2 Roadmap

Input document:

[3228](#) Roadmap update; Uma; 2007-04-18

This document is result of all the action items on the Roadmap ad hoc.

**Relevant resolution:**

**M50.42 (Roadmap snapshot):** WG2 instructs its convener to post the updated snapshot of the roadmaps (in document [N3228](#)) to the WG2 web site.

### 7.3 Request from SEI - UC Berkeley re: Minority Scripts

Input documents:

[3232](#) Letter to ISO/IEC JTC 1/SC 2 Chair; Debbie Anderson; 2007-04-08

[3233](#) Minority Script Report; Debbie Anderson; 2007-04-08

[3234](#) A letter to the SC 2/WG 2 convener to forward the letter and report to "JTC 1 chair through SC 2 chair and SC 2 - secretariat"; Debbie Anderson; 2007-04-08

Ms. Deborah Anderson: There was another letter from Dr. Mark Davis to SC2 chair endorsing the work of UC Berkeley. The letter to WG2 in document [N3232](#) gives information about Script Encoding Initiative (SEI) and is for wider distribution to SC2 and JTC1 member bodies.

**Relevant resolution:**

**M50.41 (Letter from SEI- UC Berkeley):** WG2 appreciates the valuable contributions on Minority Scripts by the SEI at UC Berkeley, and instructs the convener to forward documents [N3232](#) and [N3233](#) to SC2 and JTC1 for information.

## 8 Scripts contributions – not part of PDAM 4

Mr. Mike Ksar: There are many script contributions. We will have ad hoc groups on different topics. These are non-Han scripts, including proposals for individual characters. Ms. Deborah Anderson will chair the ad hocs on items under agenda item 8. Those who want to participate in the scripts ad hoc are welcome to participate.

Mr. Michael Everson: 16 of these documents are my contributions.

### 8.1 Avestan script

Input documents:

[3178](#) Proposal to encode the Avestan script in the BMP of the UCS SEI ; Michael Everson and Roozbeh Pournader; 2006-10-20

[3197R](#) Revised proposal to encode the Avestan script in the SMP of the UCS; UC Berkeley Script Encoding Initiative (Universal Scripts Project) -- Michael Everson and Roozbeh Pournader; 2007-01-09

Mr. Michael Everson: This proposal has taken a while to put together. It had reviews from experts in Iran and outside Iran. The controversial items were on what to do with the punctuation marks. I participated in the UTC Script committee and most of the problems have been resolved. Document [N3197](#) has one AVESTAN SEPARATION POINT and is distinguished from FULL STOP. The separation point occurs between words. There is a certain amount of variation in practice and neither middle dot nor full stop can be used for this character. It was mentioned that it could be unified with some other character but we could not find what it could be. We would like this for Amd. 5. There are 62 characters.

**Disposition:** Accept.

**Relevant resolution:**

**M50.35 (Avestan script):** With reference to document [N3197](#), WG2 accepts to encode in the standard 62 characters in code positions 10B00 to 10B35 and 10B38 to 10B3F in a new block 10B00 to 10B3F named Avestan, with their names and glyphs as shown in document [N3197](#).

### 8.2 Bopomofo characters

#### 8.2.1 Bopomofo Letter IH

**Input documents:**

- [3179](#) Proposal to encode one Bopomofo character in the UCS ; Michael Everson, H. W. Ho, and Andrew West; 2006-10-19
- [3211](#) Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16
- [3215](#) Proposed additions to ISO/IEC 10646:2003; US National Body; 2007-03-05

Mr. Andrew West: A character that looks like an upside down version of U+3113 is missing. This character was once used to denote the vowel as in zhi, chi etc. It is used in linguistic text to show the vowel usage.

Mr. Michel Suignard: The US is in favour of adding this character per document [N3215](#).

**Disposition:** Accept.

**Relevant resolution:**

**M50.19 (Bopomofo character):** With reference to documents [N3179](#) and [N3215](#) item 2, WG2 accepts to encode:

312D BOPOMOFO LETTER IH

with its glyph as shown on page 36 in document [N3211](#).

## 8.2.2 Bopomofo Letter YI

**Input document:**

- [3246](#) Proposal to encode two Bopomofo characters in UCS; TCA, Liaison Contribution; 2007-04-20

Mr. Shih-Shyeng Tseng: Document [N3246](#) has two points. Character HI is mis-spelled .. should be IH. We already discussed this. A new character 312E BOPOMOFO LETTER YI (with Horizontal bar shape) is also proposed. Evidence of use for this character is given.

**Discussion:**

- a. Mr. Michel Suignard: The current use of the YI character is fuzzy -- the direction of the character at code position U+3127 can be both ways - horizontal or vertical. To some degree either we should keep the dual behavior of the current one as is, or introduce two new characters one horizontal and one vertical.
- b. Mr. Shih-Shyeng Tseng: The horizontal text uses the horizontal bar, and vertical style uses the vertical bar. In Taiwan the use of vertical bar with official Bopomofo text should be vertical.
- c. Mr. Andrew West: The character at U+3127 has two different presentation forms; the usage is mapping to two different legacy character sets. The PRC uses the Vertical form and Taiwan uses the Horizontal. Encoding Horizontal version separately may cause additional problems. The fonts have different forms in different font packages.
- d. Dr. Asmus Freytag: The Unicode book already had the annotation on pp 431-432 that this character U+3127 has two different forms. About 40 years ago we could have added both forms when existing implementations were not impacted. Now it is too late.
- e. Mr. Shih-Shyeng Tseng: We use both symbols depending on the document. Please look at the links shown in the proposal summary form. On line dictionary from ministry of education uses the vertical form for U+3127 (2nd char in 2nd row)
- f. Mr. Michel Suignard: If you use the right font you will get the right shape -- horizontal / vertical. An annotation that both forms are possible for these can be added.
- g. Mr. Andrew West: One would not see both forms in the same document in which case font switching will do. If you have evidence that you need both the forms within the same document we can take a look at it. If they are encoded separately it will also cause problems in searching. You will probably get only half the hits.
- h. Mr. Shih-Shyeng Tseng The second link in proposal summary form shows the horizontal usage.
- i. Mr. Michael Everson: The users are forced to use the Math Division sign instead of the U+3127 Bopomofo character because the font they had used did not have the vertical form.
- j. Mr. Peter Constable: If you use the font that contains the vertical form, the page will display the vertical form on your second link.
- k. Dr. Ken Whistler: There is about 100 years of history that these two glyphs represent the same character.
- l. Mr. Masahiro Sekiguchi: I want to know the reason why you need an additional character than the existing one?
- m. Mr. Shih-Shyeng Tseng: We need both forms.
- n. Mr. Masahiro Sekiguchi: What is your requirement?

- o. Mr. Chen Zhuang: In the last page, last line, there is an example showing education document -- both forms are shown together.
- p. Mr. Masahiro Sekiguchi: Educational documents / materials for students can have both forms. Are there any other examples where these characters can be seen to be used together in the same text?
- q. Mr. Shih-Shyeng Tseng: We need both forms.
- r. Mr. Peter Constable: You need to demonstrate that writing in horizontally only you need the horizontal / vice versa. Officially the form is perpendicular to the direction of writing. Do you have example of contrary use?
- s. Mr. Shih-Shyeng Tseng: The official version is to perpendicular to the writing direction. People write both ways.

An ad hoc met.

Dr. Ken Whistler: There are issues about rendering etc. for the two forms. The ad hoc met and the request for the additional character from TCA has been withdrawn. There is no further action needed regarding their contribution.

**Disposition:** TCA has withdrawn their request.

### **8.3 8 Arabic characters for Persian and Azerbaijani**

Input documents:

- [3180R](#) Proposal to encode eight Arabic characters for Persian and Azerbaijani in the UCS; Michael Everson, Roozbeh Pournader, and Elnaz Sarbar; 2006-10-24
- [3211](#) Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16
- [3215](#) Proposed additions to ISO/IEC 10646:2003; US National Body; 2007-03-05

Mr. Michael Everson: In document [N3180R](#), one of the characters is used in Azerbaijani. All the other 7 characters are used are in earlier Persian orthographies.

**Discussion:**

- a. Mr. Michel Suignard: Referring to document [N3215](#), the US is in favour of the proposed characters.
- b. Dr. Ken Whistler: Iran and Ireland have requested this. The US position is to support the proposal and keep it open on the question of which Amd. The editor has suggested it for Amd. 4.

**Disposition:** Accept. Document [N3211](#) (pages 7 and 8) has the details. Refer [N3215](#) item 4 top of second page; includes one combining mark; for Amd.4.

**Relevant resolution:**

**M50.15 (Old Persian and Azerbaijani characters):** With reference to documents [N3180](#) and [N3215](#) item 4, WG2 accepts to encode 8 Arabic characters at code positions 0616, and 063B to 063F in the Arabic block, and 077E and 077F in the Arabic Supplement block, with their names and glyphs as shown on pages 7 to 10 in document [N3211](#). The first one of these is a combining mark.

### **8.4 4 Qur'anic Arabic letters**

Input document:

- [3185R](#) Proposal to encode four Qur'anic Arabic characters in the UCS; Michael Everson and Roozbeh Pournader; 2006-11-01
- [3211](#) Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16

Mr. Michael Everson: There are different traditions of editing the Koran. The centre for printing and publishing the Koran in Iran has need for the four combining characters proposed in document [N3185R](#).

Mr. Michel Suignard: Referring to document [N3215](#), the US is in favour of the proposed characters.

**Disposition:** Accept. Document [N3211](#) has the details page 7. Refer to document [N3215](#) item 4; top of second page has code positions etc.

**Relevant resolution:**

**M50.14 (Qur'anic Arabic letters):** With reference to documents [N3185](#) and [N3215](#) item 4, WG2 accepts to encode the following combining characters:

0617 ARABIC SMALL HIGH ZAIN  
0618 ARABIC SMALL FATHA  
0619 ARABIC SMALL DAMMA  
061A ARABIC SMALL KASRA

with their glyphs as shown in the code chart on page 7 in document [N3211](#).

## 8.5 Egyptian Hieroglyphs

**Input documents:**

- [3181](#) Towards a Proposal to encode Egyptian Hieroglyphs in Unicode ; Michael Everson and Bob Richmond; 2006-10-29
- [3182](#) Sources for the encoding of Egyptian Hieroglyphs ; Michael Everson; 2006-10-29
- [3183](#) Report on progress made at the Oxford meeting of Egyptologists; Michael Everson and Bob Richmond; 2006-10-29
- [3237](#) Proposal to encode Egyptian Hieroglyphs in the SMP of the UCS; UC Berkeley Script Encoding Initiative; Michael Everson and Bob Richmond; 2007-04-10

Mr. Michael Everson: Document [N3237](#) is the main document. WG2 has seen the earlier proposal about 10 years ago. The direction is to follow Gardiner's (*Sir Alan Henderson Gardiner*) work. It took some work of Egyptologists; they met last year in a congress of Egyptologists who were interested in IT. There were some questions as to what their expectation of Gardiner's set is. The proposal is the result of the discussion with these experts. Rendering is done by higher level protocols. The right positioning of the hieroglyphs is done with that mechanism. The document describes the naming convention, the numbering system, and several pages of character properties. In due course proposal for ordering will be made for 14651. There is a format of notation database which can be used to map Gardiner's set to other sets. This will also be used for future additional hieroglyphs. The documents [N3181](#), [N3182](#) and [N3183](#) are in support of document [N3237](#) - indicating other sources etc. The proposal is to include these in Amd. 5.

**Discussion:**

- Mr. Mike Ksar: With respect to the references in Bibliography, were there any Egyptologists from Egypt participating in the congress? Are there names of these experts from Egypt? We are encoding hieroglyphs which are of interest to Egypt as well as outside.
- Mr. Michael Everson: There were Egyptian Egyptologists there. We do not have the list of attendees in the congress.
- Dr. Asmus Freytag: The report has been made that Egyptologists from all over the world participated. That should suffice.
- Mr. Michael Everson: The bibliography refers the sources -- it is mainly Gardiner's work. There are other sources from Egypt that has been translated into French and German. We all know that there could be thousands more hieroglyphs. The proposal is for a starting set. Most of the work from Egyptian Egyptologists is on the actual text.
- Mr. Reinhold Heuvelmann: There is a note on page 3 in section 6 - a comment regarding O000 notation, should this be O0001?
- Mr. Michael Everson: Clarified their usage for classes.

**Disposition:** Accept 1063 Egyptian hieroglyphs; charts and nameslist are in document [N3237](#); for Amd. 5.

**Relevant resolution:**

**M50.29 (Egyptian Hieroglyphs):** With reference to document [N3237](#), WG2 accepts to encode in the standard 1063 hieroglyphs in code positions 13000 to 13426 in a new block 13000 to 1342F named Egyptian Hieroglyphs, with their names and glyphs as shown in document [N3237](#).

## 8.6 Additional Cyrillic characters

**Input documents:**

- [3184](#) On CYRILLIC LETTER OMEGA WITH TITLO and on CYRILLIC LETTER UK; Michael Everson and David Birnbaum; 2006-10-30
- [3194R](#) Proposal to encode additional Cyrillic characters in the BMP of the UCS; UC Berkeley Script Encoding Initiative (Universal Scripts Project); 2007-01-09
- [3213](#) Proposed block allocation Cyrillic / Bamum; Asmus Freytag - contributing editor; 2007-02-23
- [3215](#) Proposed additions to ISO/IEC 10646:2003; US National Body; 2007-03-05
- [3226](#) Proposed disposition of comments PDAM 4; Michel Suignard; 2007-04-22

### ***Glyph and name changes***

Mr. Michael Everson: There is an Irish ballot comment requesting change to names and glyphs in document [N3226](#) on Amd. 4. Document [N3184](#) is covered in [N3194](#) section 2. On page 15 of document [N3194](#), the characters do not have Slavonic style names and Glyphs. For consistency with similar other characters we suggest Latin style names and glyphs as on page 15. The Irish ballot comments are simply to bring in the consistency. Our ballot comment will reverse to Accept.

#### **Discussion:**

- a. Mr. Peter Constable: A sensible idea.
- b. Dr. Ken Whistler: This is also US Ballot comment T.3. We are on record of supporting the Irish ballot comment. The suggested solution will satisfy our ballot comment.

**Disposition:** Accept glyph and name changes proposed; include in Amd. 3.

#### **Relevant resolution (for Amd 3)**

**M50.8 (Cyrillic glyph corrections):** With reference to documents [N3184](#), [N3194](#), [N3226](#) and [N3215](#) item 5, WG2 accepts to correct the glyphs for:

0460 CYRILLIC CAPITAL LETTER OMEGA  
0478 CYRILLIC CAPITAL LETTER UK  
0479 CYRILLIC SMALL LETTER UK  
047C CYRILLIC CAPITAL LETTER OMEGA WITH TITLO,  
047D CYRILLIC SMALL LETTER OMEGA WITH TITLO  
047E CYRILLIC CAPITAL LETTER OT

with their glyphs as shown in the code chart on page 9 in document [N3194](#).

Dr. Ken Whistler: Document [N3213](#) on block arrangement for Cyrillic and Coptic suggests combining Cyrillic Ext-B and C into a single block.

### ***New characters***

Mr. Michael Everson introduced the various new proposed characters and where they are needed in using Cyrillic script. Several of these are combining. A combining character (Roof) at code position 0487 .. name, glyph etc. There is a total of 106 characters. Some of these are for modern Kurdish use. Others are for older script usage.

#### **Discussion:**

- a. Mr. Reinhold Heuvelmann: Is xxF8 combining character?
- b. Mr. Michael Everson: It is proposed elsewhere in the document.
- c. Mr. Peter Constable: Are there any additional punctuations included?
- d. Mr. Michael Everson: The Tilde is new.
- e. Dr. Asmus Freytag: Every once in a while we have a situation where we have several documents on the same topic. It will be useful to take a look at all the documents.

Ballot comment T.2 from Ireland is resolved by splitting into a and b, in proposed disposition of comments document [N3226](#). a is for Amd. 4; b is for Amd. 5; later it was decided to include b also in Amd. 4. Document [N3215](#) item 7 summarizes the complete list of code positions and names.

**Disposition:** Accepted for Amd.4 with updated code positions, modified by [N3213](#) and the revised code positions are in [N3194R](#).

#### **Relevant resolution:**

**M50.11 (Cyrillic additions):** With reference to documents [N3194R](#) and [N3215](#), WG2 accepts to encode 106 additional Cyrillic characters at code positions:

0487 in Cyrillic block (1 combining mark)  
0514 to 0523 in Cyrillic Supplement block (16 characters)  
2DF6 to 2DFF in Cyrillic Extended-A block (10 characters)  
A640 to A65F, A662 to A673, A67C to A697 in Cyrillic Extended-B block (78 characters)  
2E3B in Supplemental Punctuation Block (1 character)

with their names and glyphs as shown in document [N3194R](#).

## ***8.7 Medievalist and Iranianist punctuation characters***

Input document:

[3193](#)

Proposal to add Medievalist and Iranianist punctuation characters to the UCS; Michael Everson (editor), Peter Baker, Marcus Dohnicht, António Emiliano, Odd Einar Haugen, Susana Pedro, David J. Perry, Roozbeh Pournader; 2007-01-09

[3211](#)

Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16



Mr. Michael Everson: Some characters in document [N3193](#) that the UTC was comfortable with and others they were not. There will be an ad hoc.

**Discussion:**

- a. Dr. Ken Whistler: Ad hoc agreed to go ahead with 20 out of all the proposed characters.
- b. Mr. Michel Suignard: Reference document [N3211](#) charts on pages 34 and 35, less 2E19 and 2E3B.

**Disposition:** Accept a subset of 20 characters from [N3193](#) with glyphs, names and code positions in [N3211](#).

(**Note:** It was not clear whether the ad hoc had recommended these for Amd. 4 or Amd. 5. After the relevant resolution was adopted for Amd. 5, it was discovered that the ad hoc had agreed to include this in Amd. 4 and not Amd. 5, and these were in the final charts for Amd. 4 prepared by the contributing editor. The final resolution was updated before posting with an appropriate note.)

**Relevant resolution:**

**M50.26 (Medievalist and Iranianist punctuations):** (see note 1 below): With reference to document [N3193](#), WG2 accepts to encode the following 20 characters at code positions 2E1A, 2E1B, 2E1E, 2E1F, 2E2A, 2E2C, 2E2E, 2E2F, 2E34, 2E38, and 2E40 to 2E49 in the Supplemental Punctuation block, with their names and glyphs as shown on pages 34 and 35 in document [N3211](#).

## 8.8 Archaic Sinhala numbers

**Input document:**

[3195R](#) Proposal on archaic numbers for Sinhala; Michael Everson; 2007-01-09

Mr. Michael Everson: Letters and digits proposed are numbers used in Sinhala -- not in usage over last 200 years or so. They were in UTR #2 a long time ago. At that time they were not well attested. Many Sri Lankans today don't even recognize these. We have found many other evidences of use of these historical characters. The new document replaces document [N1473](#). They were not well attested at that time along with the Sinhala repertoire. The summary proposal form is included.

Dr. Ken Whistler: The US and UTC have reviewed this contribution. We are in favour of going forward with this for inclusion in Amd. 5.

**Disposition:** Accept 20 characters, at 0DE7-0DEF, 0DF5-0DFF ; see page 3 of [N3195R](#); for Amd. 5.

**Relevant resolution:**

**M50.28 (Archaic Sinhala numbers):** With reference to document [N3195R](#), WG2 accepts to encode 20 characters in the Sinhala block at code positions 0DE7 to 0DEF and 0DF5 to 0DFF, with their names and glyphs as shown on pages 3 and 4 in document [N3195R](#).

## 8.9 29 additional Mathematical and Symbol characters

**Input documents:**

[3198](#) 29 Additional Mathematical and Symbol Characters; Asmus Freytag (Unicode), Barbara Beeton (American Mathematical Society), Patrick Ion and Murray Sargent (MathML Working Group of the W3C), Roozbeh Pournader (Sharif FarsiWeb); 2007-01-26

[3211](#) Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16

[3259R2](#) Math; Asmus Freytag; 2007-04-23

Dr. Asmus Freytag: Documents [N3198](#) and [N3250R2](#) dated 2007-01-26 have been posted replacing previous versions. It is a set of requests for additional math characters from the ongoing STIX project, as well as from the MathML WG of W3C consortium. This set of proposed characters has been endorsed by both Unicode and the US national body.

Item 2 - Invisible operators - extend the current set; for example: 3 followed by 1/2 to indicate it is an addition implied instead of multiplier etc. INVISIBLE PLUS is proposed at U+2064.

Item 3 - Two math delimiters - for use with Math text: U+27EE Left Flattened, and U+27EF Right Flattened parentheses.

Item 4 - Two arrows that were left out earlier: U+2B45 and U+2B46 Left and Rightwards Quadruple Arrows.

Item 5 - Symbol for long division similar to square root. U+27CC Long Division.

Items 6--10 are based on a review of UTR 25, contain various shaped symbols.

Item 11 - Asterisk accents: U+20F0 Combining Asterisk Above.

Item 12 - Correct the shape of Diamond operator U+22C4

Table 2 on page 14 summarizes the list; in the BMP in various blocks. There is some urgency for these characters by the user community. The request is put these in Amd. 4. Amd. 3 will be too aggressive and Amd. 5 will be too late.

**Disposition:** Accept 29 characters for Amd. 4, with names and code positions as in Table 2 and glyphs in the different sections in document [N3129](#). See document [N3211](#) for more up to date glyphs. The glyph correction in item 12 is to be processed in Amd. 3.

**Relevant resolution (for Amd. 4):**

**M50.16 (Mathematical and Symbol characters):** With reference to documents [N3198](#), WG2 accepts to encode 29 mathematical and symbol characters at code positions:

- 2064 in the General Punctuation block, with its name and glyph from pages 22 and 23 in document [N3211](#)
- 20F0 in the Combining Diacritical Marks for Symbols block, with its name and glyph from pages 24 and 25 in document [N3211](#)
- 27CC, 27EE and 27EF in the Miscellaneous Mathematical Symbols block, with their names and glyphs from pages 30 and 31 in document [N3211](#)
- 2B1B to 2B1F, 2B24 to 2B2F 2B45, 2B46, and 2B50 to 2B54 in the Miscellaneous Symbols and Arrows block, with their names and glyphs from pages 31 and 32 in document [N3211](#).

**Relevant resolution (for Amd. 3):**

**M50.7 (Math symbol glyph correction):** With reference to documents [N3198](#) item 12, WG2 accepts to correct the glyph for:

22C4 DIAMOND OPERATOR

to the shape shown in Table 1 in document [N3198](#), with a more symmetric aspect ratio.

#### *Arabic Math characters*

Dr. Asmus Freytag: Document [N3259](#) talks about an issue the Math user community has brought up about Amd. 3.

Reference FPDAM3 chart on page 60 (document [N3186](#)). The symbols in 2B30-2B44 were introduced for Arabic math allowing arrows pointing in the reverse direction. The tilde for example on 2B41 also reverses in shape. It is OK for mirroring, but does not permit indicating the reverse operation. The solution is to have Left pointing and Right pointing with Tilde and Reverse Tilde on the arrows etc. This will permit the correct mirroring and mathematical operations to be represented. Elsewhere in the standard there are other arrows with tilde. We will need the reversed versions of those as well. The missing six characters are needed for completing the set. The glyphs are straight forward and the names are straight forward. These are proposed for FDAM3.

Mr. Martin Hosken: What about the drawbacks in the fonts?

Dr. Asmus Freytag: It is a known issue and is being addressed.

**Disposition:** Accept six characters for FDAM3. Refer to document [N3259](#) for details.

**Relevant resolution:**

**M50.3 (Arabic math symbols):** With reference to document [N3259](#), WG2 accepts to encode the following 6 characters:

2B47 REVERSE TILDE OPERATOR ABOVE RIGHTWARDS ARROW  
2B48 RIGHTWARDS ARROW ABOVE REVERSE ALMOST EQUAL TO  
2B49 TILDE OPERATOR ABOVE LEFTWARDS ARROW  
2B4A LEFTWARDS ARROW ABOVE ALMOST EQUAL TO  
2B4B LEFTWARDS ARROW ABOVE REVERSE TILDE OPERATOR  
2B4C RIGHTWARDS ARROW ABOVE REVERSE TILDE OPERATOR

with their glyphs based on descriptions provided in document [N3259](#).

## **8.10 Meitei Mayek script**

Input document:

[3206](#)

Proposal for encoding the Meitei Mayek script in the BMP of the UCS; UC Berkeley Script Encoding Initiative -  
- Michael Everson; 2007-01-12

Mr. Michael Everson: Meitei is a Tibeto-Burman language spoken in Manipur and the other states bordering Myanmar. There was an old script Meitei Mayek that was abandoned in favour of Bengali in recent times. There are older and newer orthographies. The older set is larger. It is like a Brahmic script. Older orthography has conjuncts, the modern one does not. It is similar to other Indic scripts. Annotations for users of the script are proposed. Numbers are there. I have included Danda and double

Danda. Rationale has been provided based on their shapes. We have encoded Dandas for other minority scripts of India. If these are to be unified with something else I prefer it not be Devanagari Dandas -- more akin to Saurashtrian ones. Discussions have been done with Meitei scholars as well as users in Manipur who provided the fonts for it. The Government of India is aware of it, but they did not make any comments on it. UTC has discussed this a few meetings ago.

#### Discussion:

- a. Mr. Michel Suignard: We are in favour of this script, except for the Dandas, which should be unified with one of the many sets of Dandas in the standard. It is not helpful to have too many Dandas in the standard.
- b. Mr. Michael Everson: The script identity is important for the users.
- c. Dr. Umamaheswaran: There is an open question about the Danda unification. The recommendation should be to use the Devanagari Danda.
- d. Dr. Asmus Freytag: The question of Dandas should be settled as a policy once and for all. Every time we keep opening the Dandas issues.
- e. Mr. Mike Ksar: We can go without these characters.
- f. Dr. Asmus Freytag: We can take a divide and conquer approach. It is necessary to set aside some time to settle this before we move forward. I can understand Michael's proposal. Unicode's position has been having smallest number of these. It was OK for these outside of India. There is a little bit of shift in the Unicode membership. There will be many more scripts of this nature. There is a generic issue here -- some criteria have to be brought up and discussed.
- g. Mr. Mike Ksar: We are not going to set a policy here. We need a proposal to discuss and agree on before we can go forward.
- h. Dr. Ken Whistler: Either we add or not add the Dandas and go forward with including Meiti Mayek script in Amd. 5. We do not have time to discuss the overall Dandas related item at this meeting.
- i. Dr. Umamaheswaran: My preference is to leave these out.
- j. Dr. Ken Whistler: US preference is to leave these out. But we will not vote against the resolution if they are included.

An ad hoc discussed the Danda situation. Consensus was to exclude these two from the proposal going forward into Amd. 5.

**Disposition:** Accept for Amd. 5; 76 characters in the range 1C80 -- 1CCF.

#### Relevant resolution:

**M50.33 (Meitei Mayek script):** WG2 accepts to encode in the standard 76 characters in code positions 1C80 to 1CBC, 1CBF to 1CCC and 1CCF, in a new block 1C80 to 1CCF named Meitei Mayek, with their names and glyphs as shown in document [N3206](#).

### 8.11 Bamum script

#### Input documents:

- [3209](#) Preliminary proposal for encoding the Bamum script in the BMP of the UCS; Michael Everson and Charles Riley; 2007-01-27
- [3211](#) Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16
- [3213](#) Proposed block allocation Cyrillic / Bamum; Asmus Freytag - contributing editor; 2007-02-23

Mr. Michael Everson: Bamum is a syllabic script from Cameroon devised around 1900. It was partly ideographic, partly syllabic, and went through several revisions. The modern use of these is as a syllabary. The proposal is to add 88 characters in the BMP for now to support the educational effort in this area. There are about additional 500 or so characters to be researched and would be candidates in the SMP. A set of 10 characters are used both as letters and numbers. They were originally digits but now have become letters in use. The request is to include these in Amd. 5. It is difficult to deal with the user community because they don't have internet etc. The larger set has syllabics and ideographs etc. They all have the letter property.

#### Discussion:

- a. Ms. Deborah Anderson: Names for A6F0 and A6F1, should have 'combining mark' in them.
- b. Mr. Mike Ksar: It is called preliminary proposal. It is the first time WG2 sees it. Usually a first time proposal is not sent out for ballot. Is there any urgency for moving forward with it now?
- c. Mr. Michel Suignard: We have seen this proposal -- document [N3211](#) has already included these. For us it is good enough for Amd. 5. We have done similar things in the past. If there is a controversy there would be reason to wait.

- d. Dr. Asmus Freytag: The key thing is if the experts on the script in this room anticipate the need for further work before the proposal can be balloted. In this instance this is not the case. We did not see it in UTC review either. The requirement for two technical ballots is the minimum we need. It should go into an amendment and will attract the comments from national bodies.
- e. Mr. Mike Ksar: I want WG2 members to be aware of what is happening. I don't want this to set a precedent.
- f. Dr. Umamaheswaran: There is an existing user community for the script. The proposal seems to be quite well done and I would support us moving ahead and including it in Amd. 5.
- g. Ms. Deborah Anderson: As a liaison from SEI I am comfortable with us going ahead with this at this time. It is a modern set unlike the larger set that would need quite a bit of more work.

**Disposition:** Accept 88 characters in the range A680-A6DF; refer to document [N3211](#) pp 39-40; for glyphs and corrected names; for Amd. 5.

**Relevant resolution:**

**M50.32 (Bamum script):** With reference to document [N3209](#), WG2 accepts to encode in the standard 88 characters in code positions A6A0 to A6F7 in a new block A6A0 to A6FF named Bamum, with their names and glyphs as shown in document [N3265](#). Two of these characters are combining marks.

### 8.12 Tai Viet script

**Input documents:**

- [3220](#) Expert contribution describing the Tay Viet script & completed [N3002](#) Proposal Summary Form; Jim Brase, SIL International; 2007-03-20
- [3221](#) Vietnam NB comments on Tay Viet script; Vietnam NB; 2007-03-21

Mr. NgoTrung: Document [N3221](#) is the letter from the Vietnamese national body in support of document [N3220](#).

Mr. Martin Hosken: Mr. Jim Brase has prepared document [N3220](#). The most controversial item in the discussion is visual order versus logical order within this proposal. The Unicode consortium has seen this, and the document was discussed there and they seem to be happy with it. The rationale for choice of the model is in the document.

**Discussion:**

- a. Dr. Ken Whistler: The US national body is in favour of encoding it and agrees with the approach taken in the document.
- b. Mr. Martin Hosken: There are empty positions in code charts. Many more characters are under investigation and awaiting further research. These additions are not ready. The editor has the font.

**Disposition:** Accept; in a new block Tai Viet - AA80 to AADF; 72 characters; with names, glyphs and code positions from document [N3213](#); for Amd. 5.

(**Note:** The exact number of characters is 73. Correction to the relevant resolution was made after the resolution was adopted and checking the final charts prepared by the contributing editor.)

**Relevant resolution:**

**M50.31 (Tai Viet script):** With reference to documents [N3220](#) and [N3221](#), WG2 accepts to encode in the standard 73 characters in code positions AA80 to AAC3 and AADB to AADF in a new block AA80 to AADF named Tai Viet, with their names and glyphs as shown in document [N3220](#). Several of the characters are combining marks.

### 8.13 Additions for Coptic and Latin

**Input document:**

- [3222](#) Proposal to add additional characters for Coptic and Latin to the UCS; Michael Everson, Stephen Emmel (Universität Münster), Antti Marjanen (University of Helsinki), Ismo Dunderberg (University of Helsinki), John Baines (Oxford University), Susana Pedro (Universidade Lusófona de Humanidades e Tecnologia), António Emiliano (Universidade Nova de Lisboa); 2007-03-15

Mr. Michael Everson: International Association of Copticists are working on a font to be made available to all Coptic users by January coming year. There are several Cryptogrammic letters in Coptic; some are already encoded in the standard. Two letters used in Coptic are proposed - Combining Right -Joining and Left-Joining Macrons. When the proposal was written in March, two of these characters in this document are more urgent than the others. On page 3 of the document [N3222](#), under item 4, there are letters shown with macrons on top. Some macrons go over multiple letters -- with a piece starting halfway on the first and joining with the macron on the next letter. The last one ends also halfway over the last letter. The combining Overline in the standard can be used for the macron usage over the middle line. It has different properties -- usually macron sits over the capital letter or the small letter as the case

may be. However in Coptic case, the use of the macron in Coptic stays at the uppercase letter position. Implementers have used combining Overline. There are other combining half marks in the standard already - like combining tilde left half. We propose combining macron left half, combining macron right half and combining conjoining macron -- in the earliest ballot possible. The suggested code positions are FE24, FE25 and FE26.

**Discussion:**

- a. Dr. Asmus Freytag: I have read some email discussions. A number of alternatives have been raised. None of them had the attraction of this proposal. The simplicity, attractiveness etc. of this proposal is better.
- b. Mr. Michel Suignard: Given that there is some controversy I prefer to give the national bodies a chance to review these and place this in Amd. 4 instead of in Amd. 3.
- c. Dr. Ken Whistler: Neither the US national body nor the UTC has reviewed this proposal. Amd. 3 is out of the question. I think Amd. 4 will be acceptable giving the national bodies some time to review, in view of the urgency of the Coptic implementers.
- d. Mr. Peter Constable: I would like to ask that the document be updated to include the proposal for the third one.
- e. Mr. Michael Everson: There are also other characters in the proposal we would like to add in a future amendment, but have a lower level of urgency.
- f. Dr. Ken Whistler: There is consensus to add seven additional characters for a future amendment. We should include three combining characters FE24, FE25 and FE26 in the revised document, in Amd. 4. We need seven more characters in the Coptic block 2CEB to 2CF1 from document [N3222](#) to be included in a future amendment.
- g. Mr. Peter Constable: Why the urgency?
- h. Mr. Michael Everson: the Coptic implementation does not quite work without these combining marks. The other seven are less urgent.

**Disposition:** Accept 3 FE24, FE25 and FE26 for Amd. 4. Accept remaining 7 for future Amd. 5.

**Relevant resolution - for Amendment 4:**

**M50.22 (3 Combining Macrons):** With reference to document [N3222](#), WG2 accepts to encode the following 3 combining marks:

FE24 COMBINING MACRON LEFT HALF  
FE25 COMBINING MACRON RIGHT HALF  
FE26 COMBINING CONJOINING MACRON

in the Combining Half Marks block, with their glyphs as shown in document [N3222](#).

**Note:** The documents referenced in the above resolution should be [N3222](#) instead of [N3235R](#) (an error in the adopted resolution M50.22 in document [N3254](#)).

**Relevant resolution - for Amendment 5:**

**M50.30 (Coptic additions):** With reference to documents [N3222](#), WG2 accepts to encode in the standard 7 characters in code positions 2CEB to 2CF1 in the Coptic block, with their names and glyphs as shown in document [N3222](#). Three of these characters are combining marks.

## 8.14 Lanna script

**Input documents:**

<a href="#">3207</a>	Revised proposal for encoding the Lanna script in the BMP of the UCS; Michael Everson, Martin Hosken, and Peter Constable; 2007-01-30
<a href="#">3215</a>	Proposed additions to ISO/IEC 10646:2003; US National Body; 2007-03-05
<a href="#">3226</a>	Proposed disposition of comments PDAM 4; Michel Suignard; 2007-04-22
<a href="#">3238</a>	Proposal on encoding Old Tai Lue (Lanna); China; 2007-04-03
<a href="#">3239</a>	Response to <a href="#">N3238</a> – encoding Old Tai Lue; UK and Irish NB; 2007-04-11

The Lanna-related ballot comments and proposed disposition of comments were reviewed.

*China T.1 a* - The ballot comment itself does not propose an alternate name.

Mr. Michel Suignard: We can add a note to the effect -- this script is also known as xxxx - like what we did for Tai Lue. See note under figure 4 in Amd. 1 to the standard.

**Discussion:**

- a. Mr. Peter Constable: Where the alternate name is recorded? It is in the Unicode standard.
- b. Mr. Chen Zhuang: If a new name is accepted by all concerned parties then we prefer to use that name instead of a note like what we did for Tai Lue.
- c. Mr. Mike Ksar: If this meeting decides to keep Lanna - will you have problems?

- d. Dr. Asmus Freytag: One may contrast that this is xx, yy etc. but not zz. It may answer the problem.
- e. Mr. Michael Everson: The new Tai Lue script is used in China .. the old Tai Lue is older script in China. The script itself has many more users outside of China. The Tai Lue people use a small subset in China and the usage in other countries use different subsets.
- f. Mr. Martin Hosken: The script is used for Tai Lue, Modern Thai and Khün. For good reasons, each user group thinks that it is their script. In coming up with a name for the script, we came up with a name as language neutral as possible.
- g. Mr. Peter Constable: My understanding of the history is that the script became in usage about 800 years ago, in a number of regions, under a principality -- that came to be known as Lanna. It is associated with the ruler over that entire region when the script became established.
- h. Mr. Chen Zhuang: Lanna is not used in North Thailand. Some educational books call these Mong or Dai.
- i. Mr. Martin Hosken: The language is known as Gamueng in northern Thailand.
- j. Dr. Asmus Freytag: If it is needed to be completely neutral we can call it "South East Asian script 15", and be done with it.
- k. Mr. Peter Constable: My perception is that term Lanna covers the entire region -- if it is not perceived that way in China, I don't know if I can find a term that is language neutral that would be acceptable to everyone. We can call it Old Tai Lue and give Lanna as alternate or vice versa.
- l. Mr. Anssi Yli-Jyrä: Is it possible to come up with a name referencing both kinds of names?
- m. Mr. Michael Everson: Procedurally we have a script that has passed in the ballot. The Chinese comment does not have a proposed alternative. Old Tai Lue is language specific. Lanna -- meaning million rice fields -- is neutral name for the script.
- n. Mr. Mike Ksar: I have a suggestion that may be useful. I think China would like to have their name appear in the standard - not as a note. Suggestion would be to change only the block name to something like 'Lanna-Tai Lue' and not all the character names.
- o. Mr. Martin Hosken: Are there limits on block names?
- p. Dr. Asmus Freytag: There are practical limits -- and there are limits on what can be put in a block name. The hyphen is kind of OK. In the Unicode context the block name is considered ignoring hyphens, cases etc.
- q. Dr. Ken Whistler: We cannot include a block name in parenthesis. We have included parenthetical notes in clause A.2.1. The US national body would object to a change in the normative block name. We will not object to adding annotations in clause A.2.1 in the standard.
- r. Mr. Chen Zhuang: We can go along with the US proposal.

**Disposition:**

Add parenthetical entry (Old Tai Lue) to the blocks list in clause A.2.1, in Amd. 4.  
Add a Note indicating that the parenthetical entry is not part of the block name.

*China T.1b*

There was no contribution. China withdrew the comment.

*Ireland T.1 - Requests 2 more characters*

1A2D - LANNA LETTER KHUN HIGH CHA with glyph as in document [N3207](#), and move down the rest down by one code position;

1AAD - LANNA SIGN CAANG

UK T.1.3 also is on the same two characters -- name change KHUN to KHUEN.

**Disposition:** Accept with name change to KHUEN.

*UK T.1.1* - Proposes minor name changes to several Lanna code positions; some changing KHUN to KHUEN; others removing HIGH etc.

**Disposition:** Accept.

*UK T.1.2* - Moving code positions per document [N3207](#).

**Disposition:** Accept.

*UK T.1.3* - resolved with Irish comment T.1 above.

**Disposition:** Accept.

*UK T.1.4* - Request for removing two 1A65, 1A66 from current amendment. Need more study; also US comment.

**Disposition:** Accept.

*UK T.1.5* - Moving up code positions filling the gap left by removed characters.

**Disposition:** Accept.

Details are in disposition of comments document. Net changes from above discussion - 2 characters were removed and 2 new ones were added; and several character name changes; addition of a parenthetical notation (Old Ti Lue) for Lanna in the block names list, with revised names and code positions.

Mr. Michael Everson: Document N3207 shows the table incorporating all the modifications from Ireland. There are two charts in two different font styles. The one in the amendment is as in Northern Thai font. The consensus is to use the Khuen font style on page 16 of document N3207R.

**Relevant resolution:**

**M50.10 (Lanna - changes):** WG2 notes the following significant changes to Lanna script in Amd. 4 given in document N3261:

- a. Remove Vowel Signs AM and TALL AM at 1A65 and 1A66, and move up characters at 1A67 to 1A7D by two code positions to 1A65 to 1A7B
- b. Move down characters at 1A29 to 1A5E by one code position to 1A2A to 1A5F per document N3207
- c. Add two characters:  
1A29 LANNA LETTER KHUEN HIGH CHA  
1AAD LANNA SIGN CAANG
- d. Use the Khün font style for the code chart (shown in page 16 of document N3207), and
- e. Change the names for several characters as given in document N3261
- f. Add (Old Tai Lue) to the block name Lanna in A.2.2

Add clarification that parenthetical items are not part of the block names

Mr. Chen Zhuang: What about the other characters that we need?

Dr. Ken Whistler: Ad hoc it and get these prepared at this meeting, or have it brought up as ballot comments on FPDAM4.

Mr. Martin Hosken: There are two documents N3238 from China and document N3239 is a response to that proposal.

An ad hoc met on Lanna - chaired by Ms. Deborah Anderson.

Ms. Deborah Anderson: It was decided that Chinese delegates meet with Mr. Martin Hosken and other interested experts. China is requested to add comments in FPDAM4 ballot. There is no impact at this meeting for current amendments.

### **8.15 Characters for Vedic Sanskrit**

Input document:

[3235R](#) Proposal to encode characters for Vedic Sanskrit in the BMP of the UCS; Michael Everson and Peter Scharf (editors), Michel Angot, R. Chandrashekar, Malcolm Hyman, Susan Rosenfield, B. V. Venkatakrishna Sastry, Michael Witzel; 2007-04-13

Mr. Peter Scharf: We are introducing a number of characters for Vedic Sanskrit. Some characters are as extensions to Devanagari. The document shows the characters that are already encoded. The Telugu sign that was previously added is proposed to be withdrawn and added later as part of this proposal. Udatta high pitch, Anudatta low pitch and Svarita high to low, were the original concept. These have gone through transformations. There are three types of Svaritas. - described in section 4 of the document. There is also Kashmiri independent Svarita shown in section 3. Combining digits and letters representing the various tones are listed in section 4.1, including showing where they are used. Combining digit 8 is included for completion of the set even though no evidence of its use in Samagana has been found yet. Section 4.2 lists combining diacritics used in Samavedic tradition. Section 5.1 lists combining diacritics above and below letters used in Yajurvedic tradition. Many of these marks have multiple uses in various traditions. Section 5.2 lists dots below diacritics. Section 6 shows Atharvavedic Svarita. Section 7 shows different Visarga diacritics / marks. Section 8 shows different Anusvaras - several are combining others are not. Section 9 proposes additional characters for Devanagari. Prishthamatra - looks like the

combining AA, but is placed before the consonant. This permits representing older orthographies to be accurately represented. Pushpa is already encoded in the standard  
Section 11 -- additions for Bengali and other scripts similar to the superscripted numerals for Devanagari are foreseen in the future; they are not proposed for now.

Additions are also proposed for Oriya and Malayalam scripts to be able to write Vedic Sanskrit using these scripts. A proposal summary form is attached. The charts on pages 30-41 show the proposed additions in the different blocks.

(Section 12 - does not exist in the revised proposal.)

#### Discussion:

- a. Mr. Mike Ksar: This is the first submission to WG2. There are 65 proposed characters. Are experts from the various scripts that are identified here aware of this proposal? Have they been contacted?
- b. Mr. Peter Scharf: We had a workshop in Brown University where people knowledgeable in the field were present. People from the various national bodies have opportunity to comment.
- c. Mr. Peter Constable: Under section 2 - what about the Telugu sign? Ireland has requested and WG2 accepted its removal from ballot for now. The character Udatta - has name Vedic Tone Svarita.
- d. Mr. Peter Scharf: The vertical tone above has the high tone. Sometimes it indicates Udatta and sometimes it is Svarita. The name as Udatta is partially correct -- use as Svarita is more appropriate. Probably vertical stroke above would be more appropriate. There is an annotation in Unicode.
- e. Mr. Peter Constable: There have been proposals from Government of India on Vedic to UTC. I would anticipate there will be some interaction with that proposal. Have you contacted them and compared their proposal. How complete is it?
- f. Mr. Michael Everson: I have copied the Government of India representative on this proposal. There are some characters here that they have not asked for. Their proposal also had a long list of characters which may be considered as glyph variants.
- g. Dr. Ken Whistler: There is a degree of overlap on the set of accent marks.
- h. Mr. Peter Scharf: It covers all known accents of Rig, Yajur and Sama Vedas. Some traditions are not included because it would require inclusion of the entire script superscripted.
- i. Mr. Michael Everson: If there are any accent marks in the Government of India document, it would be because we could not find attestations for them. We should continue to communicate with them. If they surface additional material then we can consider those.
- j. Mr. Peter Constable: The combining Devanagari VI -- how many others of that nature -- etc
- k. Mr. Peter Scharf: The tradition that uses the entire alphabet superscripted exists in handwritten form in Grantha script. I have not seen these printed in Devanagari script. The vi is also used in other traditions.
- l. Ms. Deborah Anderson: This is the first time the proposal is introduced in WG2. We need input from other script experts also.
- m. Mr. Mike Ksar: We would like to expose this to other national bodies and experts to get more feedback.
- n. Dr. Ken Whistler: The US national body specifically will have no problems with the three Oriya and Malayalam characters on ballot now because they are well understood filling positions structurally. These do not have the same potential problem as the others. We will not object to these being for future amendment also. Note that we do have some other Malayalam characters in Amd. 4. These are comparable to others we have encoded, having found attestations we can go forward on these. The Telugu character we removed is part of a larger collection of characters -- need more review.

**Disposition:** Accept 3 Oriya and 1 Malayalam character from document N3235 pages 32 -35. For Amd. 4.



**Relevant resolution:**

**M50.23 (Indic characters for Vedic):** With reference to document [N3235R](#), WG2 accepts to encode the following 4 combining marks:

0B44 ORIYA VOWEL SIGN VOCALIC RR  
0B62 ORIYA VOWEL SIGN VOCALIC L  
0B63 ORIYA VOWEL SIGN VOCALIC LL  
in the Oriya block, and,  
0D63 MALAYALAM VOWEL SIGN VOCALIC LL  
in the Malayalam block

with their glyphs as shown in document [N3264](#).

**Action item:**

National bodies and liaison to review and feedback on [N3235R](#).

### **8.16 Anatolian Hieroglyphs**

**Input document:**

[3236](#) Proposal to encode Anatolian Hieroglyphs in the SMP of the UCS; UC Berkeley Script Encoding Initiative; Michael Everson; 2007-04-09

Mr. Michael Everson: These are used for Indo European languages. It is not as large a character set as Egyptian hieroglyphs. During the decipherment process many of these have been given catalog numbers, and many are now identified. The names list contains cross references for normalization etc. and also for transliteration purposes. This is the first time WG2 has seen this contribution. We are in touch with many of the experts, none of whom are Anatolians who are now extinct. There are a few outstanding issues. One more character was added and this document will be updated to reflect that.

**Discussion:**

- a. Mr. Mike Ksar: Is there a relation between the Egyptian and Anatolian hieroglyphs? They may contain similar glyphs but have different meanings.
- b. Mr. Michael Everson: The development of the two was independent. A number of the characters have syllabic values but most are hieroglyphs. There is ONE combining character used to add productively R to other characters.

**Action item:**

National bodies and liaison to review and feedback on [N3236](#).

### **8.17 Parthian, Inscriptional Pahlavi and Psalter Pahlavi scripts**

**Input document:**

[3241](#) Proposal for encoding the Parthian, Inscriptional Pahlavi, and Psalter Pahlavi scripts in the BMP of the UCS; Michael Everson; 2007-04-12

Mr. Michael Everson: This document is for information at this meeting. The scripts descended from early Imperial Aramaic. Early Iranian, Middle Iranian and Late Iranian texts in Persia use these scripts. All are Right to Left scripts. They all have some numbers but different sets. Figure 5 shows some typographic ligatures in Parthian. Whether these are obligatory or optional is not known at this time. Inscriptional Pahlavi script is another one of these with a small user community. Psalter Pahlavi is another one; it is known from Psalms, has more digits than others, and has its own punctuation mark (looks like a snowman). Psalter has touching / shaping behavior like Arabic. Avestan has touching behavior for example. We need more research on this behavior of these scripts. Currently character names follow the names similar to related Avestan names. Suggestions have come back to give them Semitic names.

Mr. Mike Ksar: Are there experts besides WG2? I would like to suggest to contributors please send this document and similar contributions to other experts who are also outside of WG2 so that some feedback can be consolidated for WG2 meetings. Discussion can go on between meetings.

**Action item:**

National bodies and liaisons to review and feedback on [N3241](#).

### **8.18 Additional historic syllables for Vai**

**Input document:**

[3243](#) Proposal to add additional historic syllables for Vai to the UCS; Michael Everson; 2007-04-17

Mr. Michael Everson: Irish ballot comment requested the Vai block be extended by one column in anticipation of additional characters. Two of these are proposed in [N3243R](#). The proposal is to encode A62A and A62B from the characters shown in the charts on pages 6 and 7. Vai is in PDAM3 and it will go

forward in FDAM3. The user community has proposed glyph changes to several characters. If we don't fix it in FDAM3 we have to come back and ask for the changes again. One possibility is to include the new characters in FDAM3 based on evidence presented here. The other changes proposed are all glyph changes.

Mr. Michael Everson explained the changes in glyphs.

**Discussion:**

- a. Dr. Umamaheswaran: We must go ahead with the glyph changes right away in FDAM3. I do not see particular problems with the new characters with well documented evidence. We can go ahead with these also in FDAM3.
- b. Dr. Ken Whistler: I will support what Dr. Umamaheswaran just said. I am quite confident that we will not have problems including these in FDAM3 with these in the US national body. We are not going to get any more evidence or clarification better than what is presented here.

**Disposition:** Accept glyph changes in N3243R, Vai charts on pages 6 and 7; accept for encoding in FDAM3 A62A and A62B Vai Syllables Ndole Ma and Do.

**Relevant resolution:**

**M50.2 (Vai script glyph corrections and character additions):** With reference to documents N3243, WG2 accepts to correct the glyphs highlighted in the code charts shown on pages 6 and 7. WG2 also accepts to encode the following two additional Vai characters:

A62A VAI SYLLABLE NDOLE MA

A62B VAI SYLLABLE NDOLE DO

with their glyphs shown in document N3243.

### 8.19 Devanagari letter Candra A

**Input documents:**

[3226](#)

Proposed disposition of comments PDAM 4; Michel Suignard; 2007-04-22

[3249](#)

Proposal to Encode Devanagari Letter Candra A in the UCS – originally n2809; INCITS/L2 (US); Unicode Consortium; 2007-04-20

Dr. Ken Whistler: The specific letter Devanagari Candra A is used in Marathi. Details of justification are in document N3249. The feeling of the US national body was that it is important enough to be added into Amd. 3. Our request is that being well justified it be present as early as possible in the standard.

**Discussion:**

- a. Dr. Umamaheswaran: Canada supports the proposal.
- b. Mr. Peter Scharf: I would add that it is not only used only in Marathi but used throughout North Indian languages.
- c. Mr. Michael Everson: Good idea. It is well attested. There is no need for further review. I support its inclusion in Amd. 3.

**Disposition:** Accept Candra A as requested by US into FDAM 3. At 0972 glyph is in N3249 Devanagari Candra A.

**Relevant resolution:**

**M50.1 (Devanagari Candra A):** With reference to document N3249, WG2 accepts to encode:

0972 DEVANAGARI CANDRA A

in the Devanagari block, with its glyph as shown in document N3249.

### 8.20 Latin Capital Letter Sharp S

**Input document:**

[3227R](#)

Proposal to encode Latin Capital Letter Sharp S; Deutsches Institut für Normung (DIN), Technical Committee NIA-01-29-01 Coded Character Sets; 2007-03-21

Mr. Andreas Stötzner: The upper case version of sharp S existed in the literature over hundred years. The justification is provided in N3227. The case folding operation makes the case conversion to two 'ss' characters and makes the case conversion irreversible. Having this letter permits ability to preserve the case conversion to be reversible.

Mr. Paul Dettmer: There was an earlier decision not to accept this character in writing. If the name had sharp S, we could not write double SS in uppercase in passports. We were forced to use only small sharp S. In the past due to size limitations we could not accommodate the uppercase Sharp S. Now we do not have the coding space limitations and hence it should be possible. Our current passport application is very primitive; currently it is based on Typewriter. Some one has to write by hand the uppercase sharp S, scan it and include in the passport. The proposed character is for the future applications, permitting the government for example to go for online versions of passport etc.

## Discussion:

- a. Mr. Michel Suignard: This document keeps claiming that there is no case folding etc. But the case folding is brought up time and again. If the case folding is brought up in the future we would have problems.
- b. Dr. Asmus Freytag: It is always possible for some culture / country to come up with orthography reform which may be against what have done so far. But anticipating that ahead of time would put a break on such a reform. In many cases countries have backed away from long held belief about sorting etc. since too few vendors were able to support it. The current infrastructure's predominance will likely prevent the orthography rules to change regarding case folding etc. The proposal is well attested showing the use of the uppercase sharp S in minority situations. We can have concerns about case folding. We cannot rule it out of scope as minority character etc. The minority use of this character has been researched fairly well and it keeps surfacing as a character that has not been encoded. It is my impression from talking to DIN members; they do quite understand the issue faced by implementers that the standard orthography has well established case conversion. The minority that wants this character want the typographically correct versus orthographically correct. We should go forward and find a way of encoding this character and exclude it from the standard case conversion / pairing.
- c. Mr. Michel Suignard: I was concerned that the document is asking for inclusion but the suggestion as to 'the case conversion may potentially change' is the one raising the warning. It would break the existing model.
- d. Mr. Martin Hosken: The governments may come up with new rules.
- e. Mr. Mike Ksar: The encoding model we have -- if it is broken -- we should be concerned. We should be careful to add text in the standard about the casing operation.
- f. Dr. Asmus Freytag: The intent is not to come up with a change request. But there is possibility of change in the future.
- g. Mr. Michel Suignard: There are a couple of paragraphs - for example, the last paragraph on page 4 in the proposal.
- h. Mr. Mike Ksar: The concern raised by Mr. Michel Suignard has to be addressed.
- i. Mr. Andreas Stötzner: The document is raising issues that could be brought up. We could remove the paragraphs hinting at any case conversion change request that may be brought up in the future.
- j. Mr. Michael Everson: I think we should go ahead and put it in and move on.
- k. Dr. Ken Whistler: The US national body is opposed to including this in the standard.
- l. Mr. Masahiro Sekiguchi: What does Typographical Uppercase mean in the proposal.
- m. Mr. Andreas Stötzner: It is giving you the rationale and how the character is needed. Its origin is based on the bi-cameral nature of the script.
- n. Mr. Michael Everson: "is to be seen" is to be interpreted not as needed from the computer implementation etc.
- o. Mr. Alain LaBonté: Is the US opposed because of the property etc. One can find the equivalence using another sharp S with case conversion pairing etc.
- p. Dr. Ken Whistler: Introducing the new character like small sharp s to do the case pairing etc. will only cause more problems. I am voicing the US national body position that there is a problem. Personally I cannot parse / understand the document as to what is being proposed related to case conversion.
- q. Dr. Asmus Freytag: I can help in editing this document. We can note that the US has a position and move on.
- r. Mr. Andreas Stötzner: An updated document N3227R is produced after an ad hoc between interested experts. Paragraphs relating to IDN matters and case conversion etc. were removed.
- s. Mr. Michel Suignard: The document is better than before and is less confusing.
- t. Mr. Michael Everson: I would like to see it put on a ballot.
- u. Dr. Asmus Freytag: We left some issues on the table and they have been dealt with and we should move forward.

**Disposition:** Accept. The code position is 1E9E; glyph in N3227R. LATIN CAPITAL LETTER SHARP S; for Amd. 4.

### Relevant resolution:

**M50.18 (Latin capital letter sharp S):** With reference to document N3227R, WG2 accepts to encode:  
1E9E LATIN CAPITAL LETTER SHARP S  
in the Latin Extended Additional block, with its glyph as shown in document N3227R.

## 8.21 3 Additional Tibetan characters

Input documents:

- [3240](#) Proposal to encode 3 Tibetan; China NB; 2007-04-11  
[3247](#) Response on Tibetan proposal [N3240](#); Christopher J. Fynn, (individual expert contributor); 2007-04-20

Mr. Ngodrop: Due to problems related to sorting etc. we need three new characters.

- a. They are high frequency characters in modern Tibetan.
- b. If these are added to BMP, it becomes easier to describe modern Tibetan.
- c. Assists in sorting Tibetan properly.

Response by Mr. Christopher Fynn is in document [N3247](#)

### Discussion:

- a. Mr. Andrew West: This topic was discussed about 10 years ago. We had discussion in the committee at that time. Asking for these characters now goes against the model that is used today. It affects all existing implementations. The reasons provided by China are not strong enough to be able to justify the disruptions. The main reason for the Chinese proposal seems to be easier to sort. Mr. Christopher Fynn and others have implemented Tibetan collation and document [N3247](#) goes into some detail how to collate Tibetan successfully using the existing model. Tibetan and Zonka are already in CLDR and using those tables you can sort Tibetan correctly. Even if these characters were encoded they will not help sorting anyway. Once you encode the head letters you will come back with prefix letters. On the surface it is a good idea -- it is too late for the model and we believe it is not needed. Vista supports Tibetan properly using existing model.
- b. Dr. Ken Whistler: The US national body is supportive of the position from Mr. Christopher Fynn documented in [N3247](#). The US also does not agree that these characters should be encoded. We also agree that to encode these will be disruptive. If there is going to be further discussion, it should be along the lines of how to collate Tibetan correctly with existing encoding and not by adding new characters.
- c. Mr. Ngodrop: The item is not resolved yet. We need more discussion on the topic between experts.

**Disposition:** Needs further study between experts.

**Action item:** National bodies and liaison organizations to feed back.

### Relevant resolution:

**M50.40 (Addition of 3 Tibetan characters):** WG2 invites the Chinese national body to discuss and resolve the concerns raised at meeting M50 on the proposal for 3 additional Tibetan characters in document [N3240](#) by other WG2 experts.

## 8.22 Additional Hangul Jamos

Input documents:

- [3168](#) A Proposal to add new Hangul Jamo extended characters to BMP of UCS; NB Korea, KIM, Kyongsok, Dae Hyuk Ahn; 2006-09-27  
[3172](#) Add annotations for existing 5 Hangul Jamo names; NB Korea; KIM, Kyongsok; Dae Hyuk Ahn; 2006-09-27  
[3242](#) Proposal for allocation of proposed Hangul Jamos in the BMP; Irish NB; 2007-04-16  
[3257](#) Korean Hangul; Korean NB; 2007-04-23

Prof. Kyongsok Kim: We want to make three points - in document [N3257](#). A new model was proposed in document [N3113](#). We discussed that model at the last meeting. Now we are withdrawing that proposal. Document [N3168](#) included a proposal for adding 117 Old Hangul letters. There were some typos, and we have fixed those. We have requested annotations for five Hangul characters in document [3172](#).

### Discussion:

- a. Mr. Michael Everson: The Irish national body saw document [N3168](#). It did not propose any code positions. Document [N3242](#) is proposing suggested code positions. We are not opposed to adding these.
- b. Mr. Dae Hyuk Ahn: The positions in document [N3242](#) are OK with us.
- c. Dr. Ken Whistler: Code positions in document [N3242](#) have not been reviewed by the US national body. In the role as roadmap ad hoc member, personally I am OK with these. There is no formal position regarding adding these 117.
- d. Mr. Michel Suignard: There are ways of representing these by alternative means today. Encoding these could pose implementation difficulty.
- e. Dr. Ken Whistler: As an expert I think the best way to go forward would be to include these in a ballot and move forward.

- f. Dr. Umamaheswaran: We will have two technical ballots to deal with it. We can put it into Amd. 5.

**Disposition:** Takes note of withdrawal of a new model for Hangul Jamo usage.

Accept 117 Jamos; Document [N3242](#) has glyphs and code charts; annotations from [N3172](#). For Amd. 5.

**Relevant resolutions:**

**M50.34 (Hangul Jamo additions):** WG2 accepts to encode in the standard 117 Hangul Jamo characters as detailed in document [N3242](#). This includes creating two new blocks A960 to A97F named Hangul Jamo Extended-A and D7B0 to D7FF named Hangul Jamo Extended-B along with annotations proposed in document [N3172](#).

**M50.38 (Hangul Jamo usage model):** With reference to document [N3257](#), WG2 notes that national body of Republic of Korea has withdrawn their proposal for a new Hangul Jamo usage model in document [N3113](#).

### 8.23 Proposed additions to ISO/IEC 10646:2003

Input documents:

- [3215](#) Proposed additions to ISO/IEC 10646:2003; US National Body; 2007-03-05  
[3216](#) Proposal to Encode Additional Orthographic and Modifier Characters; Peter G. Constable, Lorna A. Pries ; 2006-10-20  
[3217](#) Proposal to encode Samaritan Text symbol; [Canada (SCC)], Society of Biblical Literature (USA), Deutsche Bibelgesellschaft (Germany); 2006-10-20  
[3218](#) Proposal to Add Additional Ancient Roman Characters to UCS; David Perry; 2006-08-01  
[3219](#) Glyph corrections for U+027F and U+0285; Lorna A Priest and Greg Aumann; 2007-01-12

#### 8.23.1 Additional Orthographic and Modifier Characters

Input document:

- [3216](#) Proposal to Encode Additional Orthographic and Modifier Characters; Peter G. Constable, Lorna A. Pries ; 2006-10-20

Mr. Peter Constable: Four characters are proposed; used in Latin orthographies in various parts of the world for minority languages. The Colon and Equals are designed to be distinct from the ASCII punctuation characters. The SALTILLO - the name is adopted from its usage in languages in Mexico - look like Apostrophes but are cased pair. Examples of use in languages are provided in the document. Saltillo is used in about 80 languages in Mexico and total about 134 worldwide.

**Discussion:**

- Mr. Michael Everson: This is one of the worst orthographic designs.
- Dr. Ken Whistler: The US did review these and is in favour of including in the current amendment. Amd. 4 is fine - nothing to be gained by pushing this into Amd. 5.
- Mr. Masahiro Sekiguchi: Question on usage of such alternate ASCII characters -- will have compatibility problems.
- Mr. Peter Constable: Appendix A has current 8-bit code pages -- they show distinct coding for Apostrophe and Saltillo.
- Mr. Masahiro Sekiguchi: Are they using Unicode? If so what code positions?
- Mr. Peter Constable: Users cannot use Unicode at this time. They are using 8-bit code positions.

**Disposition:** Accept. Four Characters from document [N3216](#) in Amd. 4. document [N3211](#) shows names, and glyphs; at A789-A78C.

**Relevant resolution:**

**M50.21 (Orthographic and Modifier characters):** With reference to document [N3216](#), WG2 accepts to encode the following 4 combining marks:

A789 MODIFIER LETTER COLON  
A78A MODIFIER LETTER SHORT EQUALS SIGN  
A78B LATIN CAPITAL LETTER SALTILLO  
A78C LATIN SMALL LETTER SALTILLO

in the Latin Extended-D block, with their glyphs as shown on page 41 in document [N3211](#).

#### 8.23.2 Symbol for Samaritan Source

Input document:

- [3211](#) Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16  
[3217](#) Proposal to encode Samaritan Text symbol; [Canada (SCC)], Society of Biblical Literature (USA), Deutsche Bibelgesellschaft (Germany); 2006-10-20

**Disposition:** Accept 214F SYMBOL FOR SAMARITAN SOURCE for Amd. 4. Refer to document [N3211](#) page 26 for correct glyph.

**Relevant resolution:**

**M50.17 (Symbol for Samaritan Source):** With reference to document [N3217](#), WG2 accepts to encode: 214F SYMBOL FOR SAMARITAN SOURCE in the Letterlike Symbols block, with its glyph as shown on page 25 in document [N3211](#).

### 8.23.3 Glyph corrections - 027F and 0285

**Input document:**

[3215](#) Proposed additions to ISO/IEC 10646:2003; US National Body; 2007-03-05  
[3219](#) Glyph corrections for U+027F and U+0285; Lorna A Priest and Greg Aumann; 2007-01-12

Item 5 in document [N3215](#) requests glyph corrections for 2 Latin characters at 027F and 0285 - glyphs to be positioned correctly with respect to the baseline - with Vertical alignment. These are primarily used for Chinese linguistics.

**Discussion:**

- a. Mr. Chen Zhuang: Why are the changes needed?
- b. Mr. Andrew West: The feedback from linguist in China was these characters did not look right. Font manufacturers also had identified the problem.
- c. Mr. Andreas Stötzner: The top of both characters is to touch the x-height on the top and the character drops lower.
- d. Mr. Mike Ksar: The rationale is in document [N3219](#).
- e. Mr. Chen Zhuang: Chinese experts have confirmed; we are OK with the corrections.

(Cyrillic glyph corrections, which are also part of item 5 from the US were dealt with in document [N3149](#)).

**Disposition:** Editorial - Include in FDAM3.'

**Relevant resolution:**

**M50.5 (Phonetic characters glyph correction):** With reference to documents [N3219](#) and [N3215](#) item 5, WG2 accepts to correct the glyphs for:  
027F LATIN SMALL LETTER REVERSED R WITH FISHHOOK, and  
0285 LATIN SMALL LETTER SQUAT REVERSED ESH  
by adjusting the vertical alignment of these glyphs as shown in document [N3219](#).

### 8.23.4 Addition of 11 Ancient Roman Characters

**Input document:**

[3211](#) Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16  
[3218](#) Proposal to Add Additional Ancient Roman Characters to UCS; David Perry; 2006-08-01

Mr. Michel Suignard: Document [N3218](#) gives the rationale - it is result of continuing investigation of more Latin characters for Ancient Roman. The list of code positions and names of those accepted by the US is in document [N3215](#) item 6. Glyphs are in document [N3218](#). Characters that are not accepted are for further investigation.

**Disposition:** Reference document [N3218](#) and document [N3215](#) item 6; for Amd. 4.

**Relevant resolution:**

**M50.24 (Ancient Roman characters):** With reference to documents [N3218](#) and [N3215](#) item 6, WG2 accepts to encode 11 additional characters at code positions:  
2185 to 2188 in the Number Forms block,  
2E19 in the Supplemental Punctuation block,  
A7FB to A7FF in the Latin Extended-D block, and  
1019B in the Ancient Symbols block,  
with their names and glyphs as shown in document [N3211](#).

## 8.24 Orkhon script

**Input document:**

[3258](#) Orkhon; Silamu; 2007-04-23

Prof. Silamu: This document replaces our earlier proposal. We received feedback from WG2 experts on our previous proposal. The document is revised. The character set, names list and code positions have been revised. We had more discussions with other interested experts at this

meeting. The result is a common code table; 40 characters, and the variants come close to 70 as shown on page 5. The revised code positions are 10C00 to 10C3F. There are also proposed changes to names.

**Discussion:**

- a. Mr. Michael Everson: We had several ad hocs. There are significant variants based on geographical distribution. Many of the characters are the same. The variations are significant enough for others. We considered the possibility of use of variation selectors. If we open the door for that there will be too many variants. The plan is to include a few characters with modified names to accommodate the variants. The plan is to come up with a revised document that will be brought up for further discussion explaining the new model. It is similar to what we did for Cyrillic.
- b. Dr. Asmus Freytag: If we can avoid the complexity of using the variation selector, we should be happier, and we can proceed on that basis.
- c. Dr. Ken Whistler: There is some similarity to how the Anglo, Germanic etc. Runes were handled. They are the same letter but significantly different in appearance that we encoded them as separate characters. The proposal made here is similar to that approach.
- d. Mr. Michel Suignard: Will the word variant-form disappear in the new proposal?
- e. Mr. Michael Everson: Yes.
- f. Mr. Andrew West: Mr. Michael Everson has said it all. Ireland and UK will work with China to come up with the revised proposal in time for the August UTC meeting and to WG2.
- g. Mr. Michel Suignard: Take note that the names proposed should follow the naming guidelines -- don't use accented letters.
- h. Mr. Mike Ksar: the new document should state clearly that it replaces the older documents [N3164](#) and [N3258](#).

**Action item:** China to prepare a revised proposal along with input from other experts.

**Relevant resolution:**

**M50.39 (Orkhon):** With reference to document [N3258](#), WG2 invites the national bodies of China, Ireland and the UK to prepare a revised proposal taking into account the discussion at meeting M50 on Orkhon, for consideration at meeting M51.

## 9 Other contributions

### 9.1 Proposal to disunify U+4039

Input document:

[3196](#) Proposal to Disunify U+4039; Andrew West and John Jenkins; 2007-01-08

Mr. Andrew West: One CJK unified ideograph character in the BMP at 4039 is a unification of two glyph shapes. According to dictionary evidence these two characters shan and jia are distinct. Evidences from old as well as modern dictionaries are provided in the document. Possible solutions are discussed in the document. Chinese writing on page 3 of the document explicitly states the characters with these components are distinct. According to the guidelines for disunification of CJK ideographs in the Principles and Procedures document the disunified character should be given a new code position at the earliest opportunity.

**Discussion:**

- a. Dr. Lu Qin: Mr. Andrew West has correctly pointed out that these should be disunified. In IRG's work the mere fact that the character appears in a dictionary is not a condition for its encoding. The unification may be incorrect. In some cases the two letters are used interchangeably. We need to get the source information.
- b. Mr. Chen Zhuang: I would like to get source information. I have to look at the glyph used in the Chinese standard.
- c. Mr. Yasuhiro Anan: The argument proposed is valid. However, we need to be careful about already encoded characters. For ex: 4096 is identified in the dictionary with some index. The TCA submission indicates that usage may be interchangeable. I would like to confirm that in usage these are distinct.
- d. Mr. Andrew West: In dictionaries they are distinct.

- e. Mr. Yasuhiro Anan: Dictionary wise they are distinct. In practice are they used interchangeably?
- f. Mr. Andrew West: In the CJK Ext. C work there were many character are being added. One of the missing characters was this one. It was pointed out to me that it had been unified.
- g. Dr. Lu Qin: They are two different characters. If you look at the second character it is a K-source. The Kanji index is pointing to one.
- h. Dr. Ken Whistler: Page 6 has the UniHan data today. Page 7 shows what the UniHan database would be after the disunification.
- i. Mr. Shih-Shyeng Tseng: TCA had submitted two characters earlier. They were unified and put in the compatibility zone at that time.
- j. Mr. Yasuhiro Anan: Why the shan form is chosen as separate character? There is another character 28E3 in the BMP where the jia form and shan form have been unified. It looks like the user can choose jia or shan forms interchangeably.

(an ad hoc met and discussed further)

- k. Dr. Lu Qin: The two characters should be disunified. According to input from TCA, source information has to be changed. We need to identify the Korea source information.
- a. Mr. Andrew West: We would like to add this character in Amd. 5.
- b. Mr. Masahiro Sekiguchi: This problem is more complicated than what we thought before the meeting. Japan needs more time to analyze the relationship with other source standards. I am not sure which direction is correct to solve the problem. It is a problem - no doubt about it.
- c. Dr. Ken Whistler: Given the evidence for the need for disunification, my preference is to resolve the problem at the earliest -- that would mean adding to Amd. 4. If there are serious source related problems and view with reference to Ext. C repertoire, I could go along with Mr. Andrew West's suggestion.
- d. Mr. Michel Suignard: By accepting this change, the font provided by the vendors for Traditional Chinese will be having an error. There is some urgency in fixing such resources. I understand the source information is needed. We do have about two months to get the source information into the Amd. 4, and the ballot time is available to review.
- e. Mr. Yasuhiro Anan: There are other possibilities of finding an example of the same nature. There are other characters with similar components. TCA input seems to indicate that they are not cognate. Dr. Lu Qin raised the issues of source references -- these are normative information.
- f. Dr. Lu Qin: The Unicode glyph is correct. The additional character source needs to be identified and the source for existing character needs to be fixed.
- g. Mr. Michel Suignard: I am referencing the font being marketed.
- h. Mr. Mike Ksar: There seems to be some urgency. If we put it into Amd. 4, we still have the ballot period to get the missing information. Putting it into Amd. 5 there will be a 6 month delay. The IRG meeting in June 2007 can review and provide verification of the information that is needed.
- i. Dr. Umamaheswaran: Errors of this nature should be fixed as early as possible. We still have 6 months of ballot time to get it right. I would support going into Amd. 4.
- j. Mr. Masahiro Sekiguchi: I would request the experts who have additional information related to sources to provide it.
- k. Mr. Andrew West: I will update the document to add more information and address some of the concerns from Japan.
- l. Dr. Lu Qin: One of the sources is DPRK - and we have been unable to get any feedback from them and I am not confident that we can get the K information in time.
- m. Mr. Michel Suignard: We can go with the current information; by default we can use the source for the new character from IRG database. The source information can be corrected.
- n. Dr. Ken Whistler: We can simply include the KP source the best we can. The Amd. 4 ballot comments can request changes or removal of the source information.



**Disposition:** Accept SHAN (as proposed on page 6), Check exact list of cross references from wording of relevant resolution.

**Relevant resolution:**

**M50.20 (Disunification of CJK ideograph 4039):** With reference to documents [N3196](#), WG2 accepts to encode the new CJK Ideograph at code position 9FC3, with the second glyph in the figure on page 1 of document [N3196](#), and the source references G\_KX0809.020, KP1-5E2B and T4-3946. WG2 also accepts to add a new source reference T6-4B7A to CJK ideograph at code position 4039 (keeping its current source references G3-5952 and KP1-5E34 and removing its source reference T4-3946).

## 9.2 Ideographic Variation Database - referencing

**Input document:**

[3256](#) Upcoming version of the Ideographic Variation Database; INCITS/L2, Unicode Technical Committee; 2007-04-20

Mr. Michel Suignard: Based on Japan's comments the reference to the Ideographic Variation Database was modified indicating that it was empty. However, now the database is NOT empty. The proposal is that the content of Amd. 3 be updated to indicate actual content of the database. I will add the exact pointer to the database. I want to ensure that Japan is aware of this change.

Dr. Ken Whistler: The database is not quite complete, even though it has gone through extensive reviews.

**Disposition:** Accept for inclusion in FDAM3 if relevant information info is available in time. Otherwise in Amd. 4.

**Relevant resolution:**

**M50.6 (IVD reference):** With reference to document [N3256](#), the project editor is instructed to update the reference to the Ideographic Variation Database at the earliest opportunity. If the relevant final information is available before FDAM3 ballot is issued it should be included in the final text of FDAM 3.

## 10 Proposed Disposition of Comments

### 10.1 FPDAM 3 ballot

**Input document:**

[3225](#) Proposed disposition of comments FPDAM 3; Michel Suignard; 2007-04-22

**Output documents:**

[3260](#) Final DoC FPDAM3; Project Editor, Michel Suignard; 2007-04-26

[3263](#) Charts Amendment 3; Michael Everson; 2007-04-27

Mr. Michel Suignard: Japan, Ireland, and US had comments  
*Ireland, Disapproval*

*Ireland T.1* - request to remove Telugu character U+0C71

Mr. Michel Suignard: The proposal is to move this character into Vedic Sanskrit extension in the future. The only concern is not to postpone this for too long.

Mr. Michael Everson: It is proposed as part of the Vedic Sanskrit proposal.

**Disposition:** Accept removing it from FPDAM3.

#### *Ireland T.2*

Two characters combining left and right joining Macrons are requested. Document [N3222](#) is referenced for rationale. See discussion and disposition under agenda item 8.13 on page 20.

#### *Ireland T.3*

Mr. Michael Everson: Ireland had postponed for further study at the last go around. There is some controversy about use of FINIAL term in the name of U+2C78. The word TAIL is typically used. The pointy ending thing is sometimes called FINIAL. From my own experience I prefer another term for the pointy ending thing.

Mr. Michel Suignard: The proposed disposition is that the term FINIAL is inappropriate and TAIL be kept unchanged.

**Discussion:**

- a. Mr. Michael Everson: The word TAIL is inappropriate. FINIAL or Bowl is preferred.
- b. Mr. Michel Suignard: A new term that is acceptable to WG2 is needed.
- c. Dr. Ken Whistler: US national body is opposed to FINIAL. If Ireland wants another term BOWL we can look at it.

An ad hoc met and discussed the comment and options and recommended the following:

**Disposition:** Replace the name with LATIN SMALL LETTER E WITH NOTCH.

*Ireland T.4*

Request to extend the block for VAI by one column.

Mr. Michael Everson explained the rationale. There are potentially more archaic characters than available empty positions in the last column.

**Disposition:** Accepted.

*Ireland Editorials:* All accepted.

Ireland reversed its ballot to Approval.

*Japan, Disapproval.*

*Japan Technical - regarding names of three collections*

Mr. Michel Suignard: The comment is a pure name issue, for new collections 286, 371 and 372, we added in Amd.3. Japan's proposal is to change 'ideographic' to 'ideographics' in the block names. Are there any objections?

**Discussion:**

- a. Mr. Michael Everson: It does not sound right to me.
- b. Mr. Michel Suignard: We have similar stuff elsewhere. It should be OK. We don't follow perfect English in ISO/IEC 10646 anyway.

**Disposition:** Accept the proposed change in block names.

Japanese vote reverses to Approval.

*UK, Approval with Editorial comments*

Accept the two comments.

*USA, Approval with comments*

*USA T.1 Name Changes*

Replace EASTERN to EXTENDED in the names in the range U+0773 to U+077D (Arabic). The input came from Iranian expert Mr. Rouzbeh Pournader. Iran is not represented here.

**Disposition:** Accept

*USA T.2 - Devanagari letter Candra A*

There is an alternate representation for this character - see document N[3249](#). There is a security issue with alternate representation.

See discussion and disposition under item 8.19 on page 26.

Dr. Ken Whistler: As a reminder to the editor, change FDAM3 to FPDAM3 in the final disposition of comments document.

**Relevant resolutions:**

**M50.4 (Disposition of FPDAM3 ballot comments):** WG2 accepts the disposition of ballot comments on FPDAM3 in document N[3260](#) and instructs its editor to prepare the final text of Amendment 3 incorporating the dispositions. The following changes are noted in particular:

- a. 0C71 TELUGU SIGN ARDHAVISARGA is removed from this amendment.
- b. 2C78 is renamed LATIN SMALL LETTER E WITH NOTCH
- c. VAI block is extended by 1 column
- d. Updates to three Collection names as provided by Japan
- e. Names of characters in the range 0773 to 077D are changed by replacing the word 'EASTERN' with 'EXTENDED' in them.
- f. 0972 DEVANAGARI CANDRA A is added (see resolutionM50.1 above on page 26)

**M50.9 (Progression of Amendment 3):** WG2 resolves to include all the items accepted for inclusion in the standard noted in resolutionsM50.4 toM50.8, into Amendment 3. WG2 instructs its project editor to forward the final text of Amendment 3 along with the disposition of comments document N[3260](#) to the SC2 secretariat for an FDAM ballot. The final set of charts and names lists are in document N[3263](#). The target starting date for FDAM is 2007-06-30.

## 10.2 PDAM 4 ballot

Input document:

[3226](#) Proposed disposition of comments PDAM 4; Michel Suignard; 2007-04-22

Output documents:

[3261](#) Final DoC PDAM4; Project Editor, Michel Suignard; 2007-04-26

[3264](#) Charts Amendment 4; Michael Everson; 2007-04-27

Mr. Michel Suignard: Comments from China, Ireland, Japan, UK and USA.

*China: Negative*

*China T.1 a and T.1b*, Related to Lanna script.

Request is to change the name Lanna to something more suitable and to add 10 more characters.

Mr. Michel Suignard: We had discussion on this in the ad hoc in M49. The ad hoc believed Lanna is the appropriate name. We had a similar issue with New Tai Lue.

See discussion and disposition related to Lanna script under item 8.14 on page 21.

*China T.2*

Mr. Michel Suignard: The request for Glyph change for 5 Mahjong tile characters. There is a similar comment from Ireland (Editorial E.2). The proposal from Ireland is acceptable except the one for PLUM.

Mr. Michael Everson: The proposed glyph is better than the one on the ballot. The one for Bamboo, the Irish comment shows the same one as on the ballot (no change). China has not proposed any Glyph.

**Disposition:** Accepted in principle -- Change the glyphs for 1F010, 1F022, 1F023, and 1F025 to those proposed in Ireland E.2 comment.

*China T.3 - CJK Ext C - 16 Glyphs in error*

Mr. Michel Suignard: We need discussion for 16 Glyphs in CJK Ext C. Table 1 in document N[3226](#) at the end shows a summary of the issues that need to be resolved, along with where the comment came from.

See discussion and disposition under item 12.2 on page 43.

*China T.4, missing last digit in documentation of code positions*

The last digit missing in some code points is a production issue. It is editorial.

**Disposition:** These will be fixed as editorial errors.

China changed its vote to Acceptance.

*Ireland: Negative*

*Ireland -T.1 on Lanna script*

See discussion and disposition related to Lanna script under item 8.14 on page 21.

*Ireland - T.2 a and b - related to Cyrillic-A*

Document N[3194](#) provides more information.

See discussion and disposition related to Cyrillic characters under item 8.6 on page 15.

*Ireland - E.1 re: Ancient Symbols*

Glyph correction for U+10194.

**Disposition:** Accepted.

As to row numbering to contain the plane number, today the row number is scoped to be within a plane. Row 01 instead of row 101 stays.

*Ireland - E.2 - covers the same set of glyphs as China T.3 comment except the Bamboo tile.*

Mr. Michel Suignard: I recommend we accept changes proposed for all except Bamboo. China has stated that they need to discuss the Plum. The Bamboo glyph stays as is -- refer to Irish comments. After an ad hoc, China had agreed to Plum glyph from Ireland.

**Disposition:** Accept all glyphs except Bamboo as proposed.

Ireland changed its vote to Acceptance.

*Japan: Negative*

*Japan - JP1 - Removing implementation Levels*

Mr. Michel Suignard: The idea was to keep only Level 3; i.e. allow all combining sequences. One can either remove all levels or just keep Level 3. Removing Level 3 was problematic -- because it was referenced by other standards such as SGML. ISO 2022 also references the Levels. In Annex N Level 3 is again mentioned. It is easier to keep one level - Level 3. Thinking more based on JP comments, the 2022 mentioning etc. is only a comment along with Esc sequences. We may be able to get rid of Levels as Normative concept, and keep it as an informative comment in the standard. For legacy reasons, we need to keep the concept of level, associated with designation. The good news is that Annex N is informative. We can accept the JP comment in principle, by moving the Level to be informative.

**Discussion:**

- a. Mr. Mike Ksar: Japan also requested a separate annex.
- b. Mr. Michel Suignard: One can always reference back editions.
- c. Dr. Umamaheswaran: ISO may not keep the old editions on the shelves.
- d. Mr. Masahiro Sekiguchi: There are several external docs that reference levels. Some users of ISO/IEC 10646 may be surprised to see there are no levels in the latest version of the standard. Japan wants to answer such situations. We can keep these in the form of Notes or an informative annex. As long as some text exists in the standard to answer any queries from outside about levels in the standard it should be OK.
- e. Mr. Michel Suignard: Section 16.2 and in UTF-8 and UTF-16 annexes we do have notes about levels. In the new draft of merged document N3230 there is a proposed text to deal with Note 2 in Annex N with levels. External standards reference dated versions of ISO/IEC 10646 when they reference levels. For ex: ISO/IEC 10646: 2003. If the 2008 version does not have Levels - it may be OK. Referencing levels in the previous editions with dated references will be OK. If you want precise wordings different from what is in the Annex N etc. I am open for these.

**Disposition:** Accept in principle. The editor is to prepare appropriate text including relevant notes.

*Japan JP 2 Ext C related*

Several problems in the CJK Ext. C charts are pointed out.  
See discussion and disposition under item 12.2 on page 43.

*Japan JP3 - convention to indicate V-source*

Mr. Michel Suignard: This comment is not quite editorial. We had a source referencing format, which is different from others. Before we had 'V single digit' and now we have 'V two digits'. We do follow whatever the national body practice is.

Mr. Masahiro Sekiguchi: Japan wants the text to reflect the national practice or convention used.

**Disposition:** Accepted in principle.

Japan changed its vote to Acceptance.

*UK Negative*

*UK T.1 :Lanna*

See discussion and disposition related to Lanna script under item 8.14 on page 21.

*UK T.2 to T.5 -- CJK Ext. C related*

Glyphs have been added in the draft disposition of comments document for easier referencing.  
See discussion and disposition under item 12.2 on page 43.

*UK E.1 and E.2 - Collection id for Unicode 5.0 versus 5.1 etc.*

Mr. Michel Suignard: Between Amd. 3 and Amd. 4 we had change in Unicode versions. The 5.1 will include both Amd. 3 and Amd. 4. The entry in Amd. 4 will be fixed by what is in Amd. 3. This is a synchronization issue.

**Disposition:** Remove it from Amd. 3 and fix it only in Amd. 4.

*UK E.3* Add annotation for Joker tile.

**Disposition:** Accept.

*UK E.4* CJK Ext C., missing last digits in code point documentation

**Disposition:** Accept.

See discussion and disposition under item 12.2 on page 43.

*US - Positive*

*US T.1 and T.2* on Lanna script

See discussion and disposition related to Lanna script under item 8.14 on page 21.

*US T.3* on Cyrillic-A

Document [N3194](#) provides more information.

See discussion and disposition related to Cyrillic characters under item 8.6 on page 15.

*US T.4 - 7* new CJK Ideographs

See discussion and disposition related to Cyrillic characters under item 12.3 on page 44.

US vote reverses to Acceptance.

#### **Relevant resolutions:**

**M50.13 (Disposition of PDAM4 ballot comments):** WG2 accepts the disposition of ballot comments on PDAM4 in document [N3261](#) and instructs its editor to prepare the final text of Amendment 4 incorporating the dispositions. The following changes are noted in particular:

- Changes to Lanna encoding as in resolutionM50.10 above
- Adjusting CJK Ext-C repertoire, source references and glyphs and moving these out for processing in a new Amendment
- Correcting the glyphs for 4 Mahjong Tiles and annotating the name for another Tile.
- Changes to Cyrillic Extension-A repertoire arising out of resolutionM50.11 above.

**M50.27 (Progression of Amendment 4):** WG2 resolves to include all the items accepted for inclusion in the standard noted in resolutionsM50.10 toM50.26 above into Amendment 4. WG2 instructs its project editor to forward the final text of Amendment 4 along with the disposition of comments document [N3261](#) to the SC2 secretariat for an FPDAM ballot. The final set of charts and names lists are in document [N3264](#). The corrected starting dates for this work item are: FPDAM 2007-06-01, and FDAM 2007-12.

## **11 Publication issues**

### **11.1 Draft results of repertoire review for FPDAM 3, FPDAM4 and future amendments**

Input document:

[3212](#) Result of Repertoire Review for FPDAM 3 of ISO/IEC 10646:2003; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16

[3211](#) Result of Repertoire Review for PDAM4 of ISO/IEC 10646:2003 and future amendments; Source: Asmus Freytag (Unicode) Status: Liaison contribution; 2007-02-16

The above documents were prepared by the contributing editors in anticipation of the discussion at this meeting based on outcome of the ballot responses, and are made available for information during the discussions at this meeting. The final charts that are output from this meeting will be different.

### **11.2 About the Code Table Format**

Input document:

[3214](#) About the Code Table Format; V.S. Umamaheswaran (recording secretary), Michel Suignard (project editor), and contributing editors; 2007-02-28

Mr. Michel Suignard: We do have two different formats for Charts and namelist for ISO/IEC 10646 and Unicode. The same contributing editor maintains both versions. This document exposes the question - do we really need to keep these two versions - and why they should be merged. The content between the two formats is very close. However, there is more information about characters that can be found in

the Unicode nameslist. It also provides a number of references to other usages of the characters --- especially in the range of symbols. Even worse, it is more work to produce both forms. It is less work for us to produce a single format for both. It will require some work on Annex P text. The annotations will now be in the nameslist. Annex P being informative, we can do these as editorial changes. The only thing that is normative is only the name. All the other information is not normative and can be added editorially. As to the CJK charts, it does not affect the ISO/IEC 10646 as much as the Unicode side. We need to think about the multicolumn charts. The objective is to come up with multiple column formats in the end for both the standards. There will be other minor changes such as removal of Rows, Planes, and Table Numbers.. Maintaining the current table numbering is a major problem. As to the timing of how to go forward, the suggestion is to move towards the changes it is proposed in the document to be for Amd. 4.

#### **Discussion:**

- a. Dr. Ken Whistler: The JTC1 ITTF style sheets / instructions / directives were for small tables within texts. The presentation could be made as a single huge table. The directive is not meaningful in the context of ISO/IEC 10646.
- b. Mr. Masahiro Sekiguchi: What is the intent of synchronization? The nameslist as well as inclusion of annotations?
- c. Mr. Michel Suignard: The intent is to include the annotations also from the Unicode side. We get this information for free in ISO/IEC 10646 because the work has to be done for Unicode anyway.. Some Annex P information may have to stay --- but these are editorial as well.
- d. Mr. Mike Ksar: To answer your question, the majority of contribution is talking about the chart generation. Beyond that some Annex material would have to be kept / revised. The intent is to include this in the working draft for the revision.
- e. Dr. Ken Whistler: One of the reasons behind this proposal is that the same tool is used for creating both ISO/IEC 10646 and Unicode charts and nameslist. One of the intent is not to have to maintain two different formats in the tool. The simplest way forward is to eliminate the switch that determines the formats etc. If the decision is made not to keep the annotations in the nameslist then the maintenance problem is still there for the tool to be able to produce one single version of the charts and the nameslist with the annotations.
- f. Mr. Masahiro Sekiguchi: I hear the short answer is YES. Will the Unicode Consortium agree to change something if WG2 is not in agreement?
- g. Dr. Ken Whistler: As maintainer of the annotations lists in Unicode, I can inform you that the Unicode consortium is open to input about annotation from all sources. Certainly WG2's opinion about annotations will be taken seriously by Unicode consortium.
- h. Mr. Michel Suignard: Once the revision FCD goes for ballot national bodies will have an opportunity on everything. Also future additions to annotations will be also open for national body comments. The key is to agree on the format first.
- i. Dr. Ken Whistler: If some national body from WG2 said -- we don't like to have cross references in the nameslist and take them all out -- it is likely that some other national bodies may say that they insist they be kept. If there are errors in annotations, or additional annotations etc. are needed then it becomes visible to Unicode as well as national bodies.
- j. Mr. Masahiro Sekiguchi: As an expert I can certainly review and provide feedback on the standard. I am confident the core members can definitely take care of problems. I am concerned how does JTC1 process dealing with conflicts? If some national body does not like the decision from Unicode Consortium how is it to be handled.
- k. Dr. Umamaheswaran: There is no change in procedures from what we are following today when we accept new characters and gets reviewed and approved by both ISO process and Unicode procedures.
- l. Dr. Ken Whistler: It is the same as having more in Annex P annotations than the current content. The same procedure is used.
- m. Mr. Mike Ksar: One of the driving forces behind is also the tool that is used by the contributing editor - Dr. Asmus Freytag. It allows keeping the maintenance of the tool simpler. If Unicode proposes some annotations, it becomes available for ISO/IEC 10646 users and if we propose something in Annex P it gets picked up in Unicode as well.
- n. Mr. Michael Everson: This is a wonderful idea -- it is preferable to have a single format for the chart / list of names. The information in the nameslist is very helpful in using the standards. For

example, there is more information about Japanese punctuation etc. for users of the standard. It is value add.

- o. Mr. Michel Suignard: That information has been provided by contributing editors to WG2 experts as information, even though the final charts for ISO/IEC 10646 did not include these in the final publications of amendments. Annotations are not on EVERY character --- it is only on those we need to provide information. It may be a good chance to expose these to a wider group of experts.
- p. Dr. Ken Whistler: The relevant nameslist file is available publicly. There is no specific action needed on behalf of the editor. I have to emphasize that the synchronization is taken very seriously on the Unicode side. The project editors produce the input to both ISO/IEC 10646 amendments as well as Unicode charts. One of the reasons for Unicode participation here is to keep the two in synch.

**Disposition:** Accept the proposed direction.

**Action item:** Editor is use the nameslist.txt from Unicode web site while preparing the next working draft - see relevant resolution **M50.37** on page 42.

**Action item:** Ad hoc on P&P to see if there is any impact on the P&P document.

### 11.3 New edition draft

Input documents:

<a href="#">3229</a>	Concept introduction (cover letter); Michel Suignard – Project Editor; 2007-04-03
<a href="#">3230</a>	New edition draft; Michel Suignard – Project Editor; 2007-04-03
<a href="#">3248</a>	Synchronization Issues for UTF-8; Ken Whistler; 2007-04-20
<a href="#">3266</a>	UCS Transformation Formats summary, non-error and error sequences; Kent Karlsson; 2007-04-25

Mr. Michel Suignard: Document [N3229](#) contains the proposal for new draft of the standard. As you know we are proposing a few amendments since 2003 edition. The more amendments we have we need to consolidate the amendments into one. What can we do to improve the standard -- in the process of creating new editions? For example, how to align it closer to Unicode? The industry is referencing both the standards more and more. It is easier to the user community if we can bring the alignments closer. There are a few roadblocks; one of them is the terminology. Terms used in the two standards have some differences. For example, Encoding Schemes, and Encoding Forms etc, in ISO/IEC 10646 have been concentrated on serialization -- leading to terms like multi-octet etc, whereas the internal representation is also important. It is worth for us to take a look at some improvements in terminology...another example is code point versus code position. They are the same but different terms. Planes, Rows etc. -- GPRC notation is pretty well obsolete -- the ranges are used nowadays instead of Planes, Rows etc. ITU has requested us to get more serialized forms such as -BE, -LE etc. This raises the reasons to distinguish between in-machine representations versus serialized forms. By having the UTF-8, UTF-16 in annexes, these are sort of treated as second class citizens. It is better to move them in to the main body. Long list of names in printed versions is better usable by having them in data files. More data files have been incorporated in the standard over time - we still have far to go. For example, the combining characters list is still maintained manually. The names list is another one; difference in formats between ISO/IEC 10646 and Unicode exists. There is no reason for the differences to be maintained. If a single format is adopted we can get rid of Annex P and have more information. My proposal is to work on a new edition towards improving the standard. Document [N3230](#) is already distributed. The version you have is .pdf version without the changes being marked. Starting document was the merged document up to Amd. 2. Amd.3 and Amd. 4 changes are shown marked up in the version I am showing. Amd. 3 changes would be shown as accepted. The changes are significant but not huge. Some parts are deeply affected, some don't change much. I will not go into details too much; perhaps a sub committee of us can review these. National bodies are requested to review all the proposed changes. Would national bodies be interested in seeing the marked up version? I would prefer the next edition to go in the proposed direction. I would like to have an ad hoc at this meeting to discuss and propose changes to the Amd. Changes in terminology would probably be the most significant from national body point of view. Should we proceed with these proposed changes in an amendment giving national bodies a chance to ballot on them? The assumption in this document is that Amd. 4 text has been accepted. The goal is to get the terminology aligned with Unicode terms. Some changes to definitions. Some new definitions have been added. The term 'character' is replaced by 'abstract character'.

## Discussion:

- a. Mr. Mike Ksar: A list of changes / replacements may be needed. One would need references to older editions.
- b. Dr. Ken Whistler: Keep character and simply say see abstract character. Being a character encoding standard you can keep it.
- c. Dr. Umamaheswaran: Is the intent to come up with a document with proposed changes and trying to get it into the current Amd. 4.
- d. Mr. Masahiro Sekiguchi: I would like to see a list of proposed changes .. item by item. The documents we have seen did not have a change history. With a version with the marked up changes we will know the details; but the rationale for each change is not clear in seeing the marked up document only. I would like to request you to prepare some background document. As to the procedure on how to deal with such changes, I would like to see it as a separate proposal. Our discussion about the 2nd edition can also go in parallel with Amd. 4. We can also assume Amd. 4 be accepted.
- e. Mr. Michel Suignard: If we assume Amd. 4 will be frozen, and have a separate track, we may find we are delaying by about 1 year. The main rationale for the proposed changes is alignment between Unicode and ISO/IEC 10646 -- to some degree, the users of both the standards is confused by having differences in terminology.
- f. Mr. Masahiro Sekiguchi: I am not objecting to the objective stated; to have same terminology between Unicode and ISO/IEC 10646 standards. For ex: CC-Data Element is used in other standards as well. If the choice is given to me I would request Unicode to change their definition to CC-Data Element etc.
- g. Mr. Michel Suignard: We have to accept the reality that Unicode is probably referenced 80% vs ISO/IEC 10646 about 20%. We have to acknowledge also the fact ISO/IEC 10646 has the machine representation vs serialization has not been differentiated. For example, the encoding forms, transformation formats etc. Coming to data-element -- it does not mean much in the context of ISO/IEC 10646 but outside it could have meaning.
- h. Mr. Masahiro Sekiguchi: CC-Data Element is only one example. Some concepts in ISO/IEC 10646 have been ambiguous. I am not objecting to the objective of synchronization. While it is important to synchronize with Unicode, synchronization with other coded character set standards in ISO is also important. Otherwise we have to change the other ISO standards also -- to keep them in alignment with ISO/IEC 10646. I would like to see a rationale document as to why for each terminology change we choose Unicode vs current ISO standard.
- i. Mr. Michel Suignard: We see all the implementers of the standards such as W3C, they are referencing Unicode definitions. If we have to be relevant to such user communities, we have to recognize that ISO/IEC 10646 is a subset of the set of definitions in Unicode. It is better if we get a CLEAN subset of Unicode into ISO/IEC 10646. That is my view on this.
- j. Dr. Ken Whistler: I would like agree with Japan's request. It is important to identify the specific changes proposed particularly in light of standards such as ISO/IEC 8859. The ISO community has been asked to move towards ISO/IEC 10646 as the reference point for coded character set terminology. What I would like to see is a set of Additional Terminology that is clearly required to cater to the growth of ISO/IEC 10646, and that it would not be a problem for other communities using ISO/IEC 10646. Others that are currently used by existing standards such as ISO/IEC 8859 we have to be more careful. We should keep the definitions; we are really not changing the definitions but changing the labels of the definition. For ex: Character to Abstract Character with the same definition. What I would like us to keep the definitions and show that particular usage of new terms used in ISO/IEC 10646. This will help other standards to continue using the current terminology in a backward compatible way.
- k. Mr. Mike Ksar: Other terms used in the past in ISO/IEC 2022, ISO/IEC 646, ISO/IEC 8859, ISO/IEC 6429 etc. that are in ISO/IEC 10646 are also valid --- the alternative terms used in ISO/IEC 10646 / Unicode can be brought out and explained. For ex: you cannot redefine the 'character' term. It is good proposal and good direction.
- l. Dr. Ken Whistler: Graphic character is one term for example that has a long history in SC2 standards. It has been extended in the Unicode standard. They have found the need to clarify use of Character etc. in Unicode. We can annotate such in ISO/IEC 10646.
- m. Dr. Umamaheswaran: To some extent we need to look at the usage of some of the terms



- n. Mr. Michael Everson: I would like to hear the national bodies point of view whether we can make it as part of another Amd. that we may come out of this meeting -- unless there is a reason for synch with other standards. National bodies can consider these changes and input on them.
- o. Mr. Mike Ksar: My thought on this would be to treat this as an amendment by itself and ballot on it in parallel.
- p. Mr. Michel Suignard: I would still try to get some agreement at the end of the meeting to progress this document. There is a lot of effort in getting this merged document out. Unless there is agreement there is no point in putting the effort. Amd. 3 and Amd. 4 the architecture part is pretty well done; there are still some open items on the repertoire point of view.
- q. Dr. Umamaheswaran: One option is to do what we did with 2003 version by treating the changes beyond Amd. 4 changes as editorial.
- r. Mr. Michel Suignard: It goes beyond just being editorial.
- s. Mr. Masahiro Sekiguchi: We can circulate as a WD of next edition ... not for balloting. At some point in Amd. 4 processing we can make the WD would be made into a final draft like FDIS etc.
- t. Dr. Ken Whistler: From US national body point of view progression of Amd. 4 should not be affected by this new proposal. This should not be treated as part of Amd. 4. I am open to other parallel processing or other means.
- u. Dr. Umamaheswaran: We can send out a WD at this meeting. FCD / FDIS etc. for next revision of ISO/IEC 10646, at the next meeting.
- v. Mr. Masahiro Sekiguch: Japan would like to have one more cycle of review in Japan. I would like to see another draft. Depends on how long we have to review.
- w. Mr. Michel Suignard: I would like to get a direction to go ahead with a Working Draft at this meeting. Seek national body comments and resolve the comments at the next meeting. We can make a decision whether to go an FCD or not at the next meeting.
- x. Dr. Ken Whistler: Japan was also requesting earlier a set of more detailed rationale for each change .. that document is important to set the ground for the FCD revision needed. Substantial changes to wordings have to be justified as well. It is not just fixing the text, but there are about 7 or 8 things about the standard is being changed. The rationale for each of the changes has to be discussed. Without such a document it is not easy to arrive at a consensus. That rewrite in the end would be useful but has to be rationalized or justified. Change to UTF-8 has to be justified. For example in document [N3248](#) there is rationale. The details are different in document [N3248](#) versus what is done here. Incorporating Annex C and D into the main body -- is also good on its own -- requires justification. Introducing terms UCS Encoding Form etc. also need justifications. On this page alone there is need for 3 distinct justifications for change in ISO/IEC 10646. Once we agree on the principles behind the justification we can arrive at a consensus on the changes proposed. I would like to see [N3229](#) rationalized to show justifications for each major category. Pretty soon after this meeting the justification docs can be distributed.
- y. Mr. Michel Suignard: One of the styles related items from JTC1 procedure is a single column mode of presentation. There may be some change in style also, and align with the JTC1 directives. Point size recommendation may not be advisable to follow -- due to reduction that happens in .pdf versions. Document [N3248](#) came in late and is on line and in the CD circulated.
- z. Dr. Ken Whistler: I wrote this document recently -- realizing there will be some discussion on the draft for new version of ISO/IEC 10646. One item that is important for implementers is UTF-8 in Annex C. The intent of UTF-8 and details of bits etc. are identical between Unicode and ISO/IEC 10646; but there is a difference between the two, in the sense that the max range between UTF-8 in ISO/IEC 10646 and Unicode exists. The Unicode has designed UTF-8 for interoperability with UTF-16 limiting the maximum range to 10FFFF. If UTF-8 is not fully interoperable as is defined in ISO/IEC 10646 with UTF-16 we do have a problem. UTF-16 and UTF-8 are the two forms that are important and widely implemented. The paper shows that there is an issue with showing error conditions. Following Annex C of ISO/IEC 10646 will lead to different error reporting with reference to code positions beyond 10FFFF. Carrying forward ISO/IEC 10646 the notion of permanent reservation of code points beyond 10FFFF, we should do the same with reference to the UTF-8 sequences dealing with beyond 10FFFF. The proposal is to leave UTF-8 in the Annex with a few changes identified. This is an example of an item which shows that there is a minor difference between Unicode and ISO/IEC 10646 but has significance in implementation. Synchronization in favour of using the Unicode limitation in this case is relevant and important in light of interoperability with UTF-16.

- aa. Mr. Mike Ksar: Is this kind of justification you are asking the editor to prepare for other changes in the Working Draft for next revision.
- bb. Dr. Ken Whistler: I have seen similar but equivalent changes in the editor's draft. I am not insisting on the precise changes that I propose be taken in as is.

**Disposition:** Editor to:

- a. Produce a rational document for each of the major category of changes in terminology introduced in the Working Draft for the next revision of ISO/IEC 10646.
- b. Circulate a marked up version of the Working Draft of the next edition showing all the significant changes needing comment from national bodies.
- c. 2003 edition will be due for 5 year review. Be ready with a new revision for it.
- d. Take note of document [N3248](#) as well.

**Relevant resolution:**

**M50.37 (WD of next edition):** With reference to documents [N3229](#), [N3230](#), [N3248](#) and [N3214](#), WG2 instructs its editor to prepare working draft of the next edition of the standard for consideration at meeting M51, incorporating the texts of Amendments 1 through Amendment 4, along with a rationale document for the proposed changes in document [N3229](#), and including the expanded names list based on the *NamesList.txt* file publicly available from the Unicode Consortium's web site.

In document [N3266](#), Mr. Kent Karlsson proposes to add some clauses regarding error handling.

Dr. Ken Whistler: What to do with error conditions is an implementation aspect. The standard does not address any error handling. The current text does not deal with error handling.

Mr. Michel Suignard: We can narrow down as to what constitutes an error condition, in dealing with unpaired characters etc.

**Disposition:**

We cannot specify what an implementation must do for error handling. Implementations do different things today.

## 12 IRG status and reports

### 12.1 IRG Meeting 27

Input document:

- [3187](#) IRG #27 Resolutions; IRG Rapporteur; Lu Qin; 2006-12-01
- [3189](#) Summary of IRG Meeting #27; IRG Rapporteur; Lu Qin; 2006-12-05

Dr. Lu Qin presented the IRG meeting 27 summary in document [N3189](#).

#### *Future meetings*

Meeting IRG 28's location has changed from Taipei to Xi'an (the date is unchanged, 2007-06-04/08).

Meeting IRG 29 will be in San Jose, CA, USA; the date is fixed.2007-11-12/16

Future meetings are being lined up. IRG needs WG2 approval for these meetings.

See relevant resolution **M50.43** on page 46.

Item 2: Consolidated all the information and ballot comments.

As to the multi column charts there is a delay on that and will be done at the next IRG meeting.

Item 3 Old Hanzi

We have accepted the principles. Second review of the input will be at next meeting.

Item 4 CJK Ext. D

Ext D contains the remainder of characters from Ext C1 work, and additional proposals from Macau.

There are also some submissions from Unicode consortium for inclusion in Ext. D.

For all future submissions IDS will be required, along with KX index, radical, First Stroke, Stroke Count, Evidence, Bitmap Image and TrueType fonts .. in electronic forms.

Item 5

For information - Errata report system will be used to manage the work of IRG. Unicode consortium requested specifically that they maintain a link to the IRG site.

Item 6

Safe characters from IICore for IDN use, is ongoing work.

Item 7

Source visual reference information is delayed. We are awaiting TrueType fonts from members. Expected to consolidate and be ready for next meeting.

Mr. Mike Ksar: Well done. Thank you.

## 12.2 CJK Ext C

Input documents:

<a href="#">3190</a>	CJK Ext. C Multi-column Code Table; IRG Rapporteur; Lu Qin; 2006-12-04
<a href="#">3191</a>	CJK Ext. C Source reference mapping table; IRG Rapporteur; Lu Qin; 2006-12-04
<a href="#">3192</a>	CJK Ext. C Source reference name table; IRG Rapporteur; Lu Qin; 2006-12-04
<a href="#">3208</a>	PDAM4; Project Editor; 2006-12-21
<a href="#">3226</a>	Proposed disposition of comments PDAM 4; Michel Suignard; 2007-04-22
<a href="#">3244</a>	Review of CJK-C Repertoire; UK NB; 2007-04-14

Dr. Lu Qin: An ad hoc met and considered the different documents. We have reviewed the first three documents on CJK Ext C. We have gotten feedback from the IRG members. There is consensus on how to go forward. Mr. Michel Suignard has compiled the images etc. Offline it should be reviewed by IRG and interested experts. Japan, TCA, UK, US, China and SC2 chair had all participated.

### Discussion:

- a. Mr. Mike Ksar: Do you have any recommendation re: Ext C in Amd. 4.
- b. Dr. Lu Qin: Do we want to continue with Ext C in Amd. 4 or do we want to delay the work till all the issues are resolved. Japan did have some comments. The general agreement is that we still would like to proceed. As to the questionable characters they should be moved out Amd.4. There was a version problem of the V font and that was part of Japanese concern. That was resolved. The UK comment was resolved.
- c. Dr. Ken Whistler: I was not involved with the details. When you say some characters are removed and you still want to go forward with Amd. 4 -- I want to be clear that Amd. 4 will go forward with only agreed upon set of characters.
- d. Mr. Michel Suignard described the draft of the summary of changes for ballot resolution.
- e. Mr. Masahiro Sekiguchi: Glyph changes - we will go closer to the original source. Not identical to what we had requested.
- f. Dr. Lu Qin: UK had asked for clarification for several characters. The ad hoc decided to remove these from Amd. 4.
- g. Dr. Ken Whistler: Mr. Michelle Suignard is doing what will be in the disposition of comments document. Another contribution which answers questions like what will be removed, what will happen to those, what will be considered by IRG etc with some rationale is needed.
- h. Mr. Mike Ksar: Document N[3267](#) will contain the IRG ad hoc group's discussions.
- i. Dr. Lu Qin: An open issue is UTC00022 versus GHC501162 for 2B151. At this time we need to have some work to do. We could remove it from current Amd. 4. We can entertain via ballot comment once the source reference and glyph misalignments are sorted out.
- j. Mr. Michel Suignard: Regarding the addition of 7 Ideographs from the US, IRG's input was not to put these into current Ext. C.
- k. Dr. Lu Qin: Seven Unicode Ideograph characters were submitted late, because of quality control issues etc. IRG decision was not to object to these additions, but outside the Ext. C collection. If you are going to ask IRG Technical editor we still need the review.
- l. Mr. Masahiro Sekiguchi: Japan is not comfortable going forward with Amd. 4. We were involved in Ext. C development and we were pushing for it strongly. At that time we had the confidence in the quality. However, after Amd. 4 ballot review, we found a lot of mistakes in the Amd. Many were editorials and we thought they can be fixed easily. But we are finding more problems after the ballot was closed. We evaluated the comments from other national bodies. So we are

reviewing the ext C again, and we found more errors. I would like to find out from other IRG members why such errors are included in Ext. C. If there is reasonable explanation for including these then we can go forward. We have NOT received reasonable explanation. So many Vietnamese sample glyphs contain errors, and they are changing the glyphs in the process. Japan wanted to know why the large number of errors were made .. we have not got satisfactory responses. Also IRG project editor had added a few characters in Ext. C; these should not have been included. There is some quality control problem. Unless we get some reasonable answers for our concerns, we are not comfortable going forward with it in Amd. 4.

- m. Dr. Lu Qin: After the number of characters are removed from Ext. C; reduces by 7 from total of 4000+.
- n. Mr. Masahiro Sekiguchi: Japan would not like to see shrinking of the tables based on removal of ideographs
- o. Mr. Mike Ksar: Are the holes to be kept?
- p. Mr. Masahiro Sekiguchi: The holes could be kept - but not to reintroduce those characters, but to put in other characters.
- q. Dr. Ken Whistler: If such a thing would be done, i.e. leaving holes etc. it would be unacceptable for the US for Amd. 4. If this is a requirement for removal and compacting we would need two ballot rounds.
- r. Mr. Andrew West: We would be opposed to leaving holes in the final table. We have seen some of Ext. C on web sites from China. If we leave holes they may be tempted to use them. By compacting it we want to ensure such implementations would change.
- s. Mr. Chen Zhuang: If it is Amd. 5, when will we start?
- t. Mr. Mike Ksar: It will be scheduled to have our ballot resolution in China meeting.
- u. Dr. Lu Qin: Would it be before the next IRG? This would mean compacting the tables.
- v. Dr. Ken Whistler: US national body will not be opposed to Amd. 4 if there are NO holes, i.e. the table is compacted. Only if there are holes left we would ask for Amd. 5.
- w. Mr. Mike Ksar: Amd. 5 means a 6-months delay.
- x. Mr. Andrew West: UK will not object to moving it to Amd. 5 giving us more time to review. Japan was indicating there are more errors.
- y. Dr. Lu Qin: The personal opinion is that I have enough evidence that another round of review will be necessary. I personally will not be opposed to moving it to Amd. 5. There are no official documents submitted on editorial errors.
- z. Mr. Shih-Shyeng Tseng: We feel it is OK to move to Amd. 5.
- aa. Mr. Chen Zhuang: Amd. 5 is OK.
- bb. Prof. Kyongsok Kim: Amd. 5 is OK.
- cc. Mr. Andrew West: Amd. 5 is OK.
- dd. Dr. Ken Whistler: The US is OK with Amd. 5.
- ee. Mr. Michel Suignard: We still need to decide what goes into the repertoire for including in Amd. 5.
- ff. Dr. Lu Qin: IRG members have agreed to remove questionable ones -- six have been removed.
- gg. Mr. Andrew West: If after the IRG meeting we get the PDAM comments we can consider them.
- hh. Document
  - ii. Dr. Lu Qin: TCA has been trying to host the IRG meeting and their effort should be acknowledged.

Mr. Mike Ksar: Thanks to Dr. Lu Qin, all the experts who participated in the discussion on Ext. C.

**Disposition:** Accept moving CJK Ext. C to Amd. 5 from Amd. 4. Document N[3267](#) will contain the IRG ad hoc group's discussions.

**Relevant resolution:**

**M50.12 (CJK Unified Ideographs Extension C):** With reference to the report of the ad hoc in document N[3267](#), and the summary table in the PDAM4 disposition of ballot comments document N[3261](#), related to CJK Ext-C, WG2 resolves to remove six ideographs from the current CJK Ext-C repertoire and move the remaining CJK Ext-C characters out of Amendment 4 and process it as part of a new Amendment 5.

### **12.3 Addition of 7 CJK Unified Ideographs**

Input document:

[3210](#) Addition of seven CJK Unified Ideographs; Eric Muller - US TAG, INCITS/L2; 2007-02-08

This document was considered by CJK ad hoc. This set is outside the CJK Ext. C set.

**Disposition:** Accept in Amd. 4.

**Relevant resolution:**

**M50.25 (CJK Unified Ideographs additions):** WG2 accepts to encode the 7 CJK Unicode-sourced ideographs, with their glyphs as shown in document N3210 at code positions 9FBC to 9FC2.

## 13 Defect reports

There were some glyph errors reported in different contributions discussed under other agenda items. The errors are to be fixed at the earliest opportunity. Most of these were accepted for inclusion in FDAM3.

## 14 Liaison & national body reports

### 14.1 Unicode Consortium

Dr. Asmus Freytag: Since the Tokyo WG2 meeting Unicode has published the Unicode 5.0 book. You have seen copies of the book here. We are very appreciative of the cooperation from WG2 to keep the ISO/IEC 10646 and Unicode in synch. We looked long and hard at the format to be used for the 5.0 edition whether this is the right way to distribute a standard and informative content. We did some market research and we did adjust the format, the size of pages etc. History of sales of 5.0 book will influence the future such publications. You can advertise the book in your own countries. Unicode has reduced prices for some kinds of institutions. Working through other book sellers the standard is available more easily outside of North America than ever before. From the standardization point of view we are in the middle of updates. Currently the Unicode Database on properties is being updated towards 5.0.1. There will be minor update to Unicode to capture Amd. 3 and Amd. 4 towards Unicode 5.1. Currently the schedule for this is not finalized. Unicode is working with people on different ways to spread the word. There was conference in Washington with international participants. There will be a conference in the fall in the West Coast of USA.

Mr. Mike Ksar: There is a link on the Unicode web site announcing reduced prices for Government Agencies, standards organizations and other such organizations. I would like to encourage people to get those because the Unicode book has lot more information for implementers much of which is out of scope for ISO/IEC 10646. You can spread the message within your countries.

Dr. Asmus Freytag: As the ad hoc groups were meeting, I have used the chance to convey our thanks from the Unicode consortium for your participation in keeping the synchronization between ISO/IEC 10646 and Unicode standards. I also want to inform you that this will be my last meeting.

Mr. Mike Ksar: I would like to thank you for your active participation and help for WG2 over the past several years. We will miss you.

Dr. Asmus Freytag: Good luck in keeping the work level.

**Relevant resolution:**

**M50.44 (Contributing Editor):** WG 2 thanks Dr. Asmus Freytag for the excellent support provided by him as a contributing editor of ISO/IEC 10646 and his valuable contributions as a WG2 expert, over the last several years. WG2 wishes him all the success in his future endeavours.

### 14.2 SEI – UC Berkeley

**Input document:**

[3231](#) Liaison report from SEI; Debbie Anderson; 2007-04-08

Ms. Deborah Anderson: If you have any groups trying to get minority scripts or ancient scripts you can contact SEI. We will be able to assist them. We have funding for the next two years and we have a long list of scripts that we are working on.

## 15 Other business

### 15.1 Web Site

Mr. Mike Ksar: I tried to keep the web site up to date. I often upload documents silently. Periodically I update the register on a monthly basis as well as prior to each meeting. I received some feedback from Dr. Umamaheswaran and Dr. Ken Whistler about some out of date information, broken links etc. and I have fixed these. The report to SC2 would be available in draft form prior to next WG2 meeting. It takes a lot of time to fix the various pages. Your feedback on making improvements would be appreciated. I will work with other officers of WG2 to see if we can have mirror image or where we can post documents,

if we can create mailing lists and some discussion forums etc. which may facilitate discussions on specific topics between meetings.

**Relevant resolution:**

**M50.45 (Appreciation to DKUUG for web site support):** 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

- Meeting 51 – 2007-09-17/21 – Hangzhou, China – with SC2 plenary and OWG-SORT.  
Mr. Mike Ksar: The location has been changed to Hangzhou from Urumqi. I had feedback from several members that Urumqi is not easily reachable. Hopefully by end of May we will have detailed logistics. We would need Visas for China. The date has not changed.  
Dr. Umamaheswaran: We can suggest that OWG-SORT be on Friday AM and SC2 on Friday PM, leaving full 4 days for WG2. Locations and dates are changed for future meetings.  
Dr. Lu Qin: HK can host at the university as long as the schedule is outside of the school terms.
- Meeting 52 – April 2008, location is changed to USA West Coast.  
Mr. Mike Ksar: The location of the meeting is changed to USA, West Coast.
- Meeting 53 – Fall 2008; tentative, Tokushima (near Osaka), Japan – with SC2 plenary
- Meeting 54 – Spring 2009 (late May) – tentative Hong Kong
- Meeting 55 – Fall 2009 – seeking host

**Relevant resolution:**

**M50.43 (Future meetings):**

WG 2 meetings:

Meeting 51 - 2007-09-17/21, Hangzhou, China (note new location); along with SC2 plenary.

Meeting 52 - 2008-05-19/23 (to be confirmed), West Coast, USA.

Meeting 53 - 2008-10, (tentative) Tokushima (near Osaka), Japan; along with SC2 plenary.

Meeting 54 - 2009-05, (tentative) Hong Kong SAR.

Meeting 55 - 2009-10 - seeking host.

IRG meetings:

IRG #28 - 2007-06-04/08, Xi'an, China (note new location)

IRG #29 - 2007-11-12/16, San Jose, California, USA (host Adobe; to be confirmed).

## 16 Closing

### 16.1 Approval of Resolutions of Meeting 50

Output documents

[3254](#)

Resolutions Meeting 50 – Frankfurt; Uma/Ksar; 2007-04-29

[3262](#)

Project subdivision proposal Amendment 5; Mike Ksar; 2007-04-29

[3265](#)

Charts Amendment 5; Michael Everson; 2007-04-27

The draft resolutions prepared by the recording secretary with the assistance of the drafting committee were discussed and adopted one by one, some with changes. The final set of resolutions is in document [N3254](#).

**Action item:** Convener and project editor to prepare a new sub division proposal for work on Amd. 5.

**Relevant resolution:**

**M50.36 (Amendment 5 – subdivision and PDAM text):** WG2 instructs its editor to prepare a project sub division proposal and PDAM text based on resolutions M50.27 to M50.35 above, and forward them to the SC2 secretariat for ballot. The proposed start dates for the progression of this work item are: PDAM 2007-05-10, FPDAM 2007-10-30, and FDAM 2008-07.

Due to the tight schedule prior to meeting M51 national bodies are encouraged to copy the project editor on their ballot comments at the same time they submit them to SC2 secretariat.

The following changes arising out of post meeting reconciliation of the charts created by the contributing editors and the adopted resolutions have been made in the final set of resolutions. The resolutions have been also repositioned and renumbered. The character count information has been updated.

**Note 1:** The ad hoc had accepted the 20 'Medievalist and Iranianist punctuations' for Amendment 4 and not for Amendment 5.

**Note 2:** The 'Tai Viet script' has 73 characters instead of 72. The first range of code positions is corrected to 'AA80 to AAC3' instead of 'AA80 to AAC2'.

## Appreciation to the Host

### Relevant resolution:

**M50.46** (Appreciation to Host): WG 2 thanks the German National Library and the German national body (DIN), and in particular Ms. Ute Schwens and Mr. Reinhold Heuvelmann, for hosting the meeting, for providing excellent meeting facilities, and their kind hospitality.

## 16.2 Adjournment

The meeting was adjourned at 12:05 on Friday, 2007-04-27.

## 17 Action Items

All the action items recorded in the minutes of the previous meetings from M25 to M46, and M49, have been either completed or dropped. Status of outstanding action items from earlier meeting M47 and M48, and new action items from the latest meeting M50 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](#))
- Meeting 49, 2006-09-25/29, Tokyo, Japan; 2006-09-25/29 (document [N3153](#))

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

Item	Assigned to / action (Reference resolutions in document <a href="#">N2954</a> , and unconfirmed minutes in document <a href="#">N2953</a> for meeting 47 - with any corrections noted in section 3 of document <a href="#">N3103</a> from meeting 48).	Status
AI-47-5	IRG Rapporteur (Dr. Lu Qin)	
	To take note of and act upon the following items.	
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. M48, M49, M50 - in progress.	

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

Item	Assigned to / action (Reference resolutions in document <a href="#">N3104</a> , and unconfirmed minutes in document <a href="#">N3103</a> for meeting 48 - with any corrections noted in section 3 of document <a href="#">N3153</a> from meeting 49).	Status
AI-48-7	US national body (Asmus Freytag)	
b.	To prepare updated Arabic Math proposal(s) based on documents <a href="#">N3085</a> to <a href="#">N3089</a> . M48, M49, M50 - in progress.	In progress

### 17.3 New action items from meeting 50, 2007-04-23/27, Frankfurt-Am-Main, Germany

Item	Assigned to / action (Reference resolutions in document <a href="#">N3254</a> , and unconfirmed minutes in document <a href="#">N3253</a> for meeting 50 - this document you are reading)	Status
AI-50-1	Recording Secretary - Dr. V.S. UMaaheswaran	
a.	To finalize the document <a href="#">N3254</a> containing the adopted meeting resolutions and send it to the convener as soon as possible.	

Item	Assigned to / action (Reference resolutions in document <a href="#">N3254</a> , and unconfirmed minutes in document <a href="#">N3253</a> for meeting 50 - this document you are reading)	Status
	b. To finalize the document <a href="#">N3253</a> containing the unconfirmed meeting minutes and send it to the convener as soon as possible.	
<b>AI-50-2</b>	<b>Convener - Mr. Mike Ksar</b>	
	<i>To take note of and act upon the following items.</i>	
	a. <i>M50.42 (Roadmap snapshot):</i> WG2 instructs its convener to post the updated snapshot of the roadmaps (in document <a href="#">N3228</a> ) to the WG2 web site.	
	b. <i>M50.41 (Letter from SEI- UC Berkeley):</i> WG2 appreciates the valuable contributions on Minority Scripts by the SEI at UC Berkeley, and instructs the convener to forward documents <a href="#">N3232</a> and <a href="#">N3233</a> to SC2 and JTC1 for information.	
	c. <i>M50.36 (Amendment 5 – subdivision and PDAM text):</i> WG2 instructs its editor to prepare a project sub division proposal and PDAM text based on resolutions M50.27 to M50.35 above, and forward them to the SC2 secretariat for ballot. The proposed start dates for the progression of this work item are: PDAM 2007-05-10, FPDAM 2007-10-30, and FDAM 2008-07. Due to the tight schedule prior to meeting M51 national bodies are encouraged to copy the project editor on their ballot comments at the same time they submit them to SC2 secretariat.	
	d. To add the following carried forward scripts to next meeting agenda, if any updates are available: Manichaeian script – document <a href="#">N2544</a> Avestan and Pahlavi script– document <a href="#">N2556</a> Dictionary Symbols – document <a href="#">N2655</a> Samaritan Pointing characters – document <a href="#">N2758</a> Babylonian Pointing characters – document <a href="#">N2759</a> Bantu Phonetic Click characters – document <a href="#">N2790</a> Invisible Letter – document <a href="#">N2822</a> Palestinian Pointing characters - document <a href="#">N2838</a> Kaithi script - document <a href="#">N3014</a> Orkhon script - document <a href="#">N3258</a> Tibetan additions - document <a href="#">3240</a> Parthian, Inscriptional Pahlavi, and Psalter Pahlavi scripts -document <a href="#">N3241</a> Anatolian Hieroglyphs - document <a href="#">N3236</a> Vedic Sanskrit - document <a href="#">N3235</a>	
<b>AI-50-3</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>	
	a. <i>M50.1 (Devanagari Candra A):</i> With reference to document <a href="#">N3249</a> , WG2 accepts to encode: 0972 DEVANAGARI CANDRA A in the Devanagari block, with its glyph as shown in document <a href="#">N3249</a> .	
	b. <i>M50.2 (Vai script glyph corrections and character additions):</i> With reference to documents <a href="#">N3243</a> , WG2 accepts to correct the glyphs highlighted in the code charts shown on pages 6 and 7. WG2 also accepts to encode the following two additional Vai characters: A62A VAI SYLLABLE NDOLE MA A62B VAI SYLLABLE NDOLE DO with their glyphs shown in document <a href="#">N3243</a> .	
	c. <i>M50.3 (Arabic math symbols):</i> With reference to document <a href="#">N3259</a> , WG2 accepts to encode the following 6 characters: 2B47 REVERSE TILDE OPERATOR ABOVE RIGHTWARDS ARROW 2B48 RIGHTWARDS ARROW ABOVE REVERSE ALMOST EQUAL TO 2B49 TILDE OPERATOR ABOVE LEFTWARDS ARROW 2B4A LEFTWARDS ARROW ABOVE ALMOST EQUAL TO 2B4B LEFTWARDS ARROW ABOVE REVERSE TILDE OPERATOR 2B4C RIGHTWARDS ARROW ABOVE REVERSE TILDE OPERATOR with their glyphs based on descriptions provided in document <a href="#">N3259</a> .	



Item	Assigned to / action (Reference resolutions in document <a href="#">N3254</a> , and unconfirmed minutes in document <a href="#">N3253</a> for meeting 50 - this document you are reading)	Status
d.	<p><i>M50.4 (Disposition of FPDAM3 ballot comments):</i> WG2 accepts the disposition of ballot comments on FPDAM3 in document <a href="#">N3260</a> and instructs its editor to prepare the final text of Amendment 3 incorporating the dispositions. The following changes are noted in particular:</p> <p>0C71 TELUGU SIGN ARDHAVISARGA is removed from this amendment.            2C78 is renamed LATIN SMALL LETTER E WITH NOTCH            VAI block is extended by 1 column            Updates to three Collection names as provided by Japan            Names of characters in the range 0773 to 077D are changed by replacing the word 'EASTERN' with 'EXTENDED' in them.            0972 DEVANAGARI CANDRA A is added (see resolutionM50.1 above)</p>	
e.	<p><i>M50.5 (Phonetic characters glyph correction):</i> With reference to documents <a href="#">N3219</a> and <a href="#">N3215</a> item 4, WG2 accepts to correct the glyphs for:            027F LATIN SMALL LETTER REVERSED R WITH FISHOOK, and            0285 LATIN SMALL LETTER SQUAT REVERSED ESH            by adjusting the vertical alignment of these glyphs as shown in document <a href="#">N3219</a>.</p>	
f.	<p><i>M50.6 (IVD reference):</i> With reference to document <a href="#">N3256</a>, the project editor is instructed to update the reference to the Ideographic Variation Database at the earliest opportunity. If the relevant final information is available before FDAM3 ballot is issued it should be included in the final text of FDAM 3.</p>	
g.	<p><i>M50.7 (Math symbol glyph correction):</i> With reference to documents <a href="#">N3198</a> item 12, WG2 accepts to correct the glyph for:            22C4 DIAMOND OPERATOR            to the shape shown in Table 1 in document <a href="#">N3198</a>, with a more symmetric aspect ratio.</p>	
h.	<p><i>M50.8 (Cyrillic glyph corrections):</i> With reference to documents <a href="#">N3184</a>, <a href="#">N3194</a>, <a href="#">N3226</a> and <a href="#">N3215</a> item 5, WG2 accepts to correct the glyphs for:            0460 CYRILLIC CAPITAL LETTER OMEGA            0478 CYRILLIC CAPITAL LETTER UK            0479 CYRILLIC SMALL LETTER UK            047C CYRILLIC CAPITAL LETTER OMEGA WITH TITLO,            047D CYRILLIC SMALL LETTER OMEGA WITH TITLO            047E CYRILLIC CAPITAL LETTER OT            with their glyphs as shown in the code chart on page 9 in document <a href="#">N3194</a>.</p>	
i.	<p><i>M50.9 (Progression of Amendment 3):</i> WG2 resolves to include all the items accepted for inclusion in the standard noted in resolutionsM50.4 toM50.8, into Amendment 3. WG2 instructs its project editor to forward the final text of Amendment 3 along with the disposition of comments document <a href="#">N3260</a> to the SC2 secretariat for an FDAM ballot. The final set of charts and names lists are in document <a href="#">N3263</a>. The target starting date for FDAM is 2007-06-30.</p>	
j.	<p><i>M50.10 (Lanna - changes):</i> WG2 notes the following significant changes to Lanna script in Amd. 4 given in document <a href="#">N3261</a>:            Remove Vowel Signs AM and TALL AM at 1A65 and 1A66, and move up characters at 1A67 to 1A7D by two code positions to 1A65 to 1A7B            Move down characters at 1A29 to 1A5E by one code position to 1A2A to 1A5F per document <a href="#">N3207</a>            Add two characters:            1A29 LANNA LETTER KHUEN HIGH CHA            1AAD LANNA SIGN CAANG            Use the Khün font style for the code chart (shown in page 16 of document <a href="#">N3207</a>), and Change the names for several characters as given in document <a href="#">N3261</a>            Add (Old Tai Lue) to the block name Lanna in A.2.2            Add clarification that parenthetical items are not part of the block names</p>	

Item	Assigned to / action (Reference resolutions in document <a href="#">N3254</a> , and unconfirmed minutes in document <a href="#">N3253</a> for meeting 50 - this document you are reading)	Status
k.	<p><i>M50.11 (Cyrillic additions):</i> With reference to documents <a href="#">N3194R</a> and <a href="#">N3215</a>, WG2 accepts to encode 106 additional Cyrillic characters at code positions:  0487 in Cyrillic block (1 combining mark)  0514 to 0523 in Cyrillic Supplement block (16 characters)  2DF6 to 2DFF in Cyrillic Extended-A block (10 characters)  A640 to A65F, A662 to A673, A67C to A697 in Cyrillic Extended-B block (78 characters)  2E3B in Supplemental Punctuation Block (1 character)  with their names and glyphs as shown in document <a href="#">N3194R</a>.</p>	
l.	<p><i>M50.12 (CJK Unified Ideographs Extension C):</i> With reference to the report of the ad hoc in document <a href="#">N3267</a>, and the summary table in the PDAM4 disposition of ballot comments document <a href="#">N3261</a>, related to CJK Ext-C, WG2 resolves to remove six ideographs from the current CJK Ext-C repertoire and move the remaining CJK Ext-C characters out of Amendment 4 and process it as part of a new Amendment 5.</p>	
m.	<p><i>M50.13 (Disposition of PDAM4 ballot comments):</i> WG2 accepts the disposition of ballot comments on PDAM4 in document <a href="#">N3261</a> and instructs its editor to prepare the final text of Amendment 4 incorporating the dispositions. The following changes are noted in particular:  Changes to Lanna encoding as in resolutionM50.10 above  Adjusting CJK Ext-C repertoire, source references and glyphs and moving these out for processing in a new Amendment  Correcting the glyphs for 4 Mahjong Tiles and annotating the name for another Tile.  Changes to Cyrillic Extension-A repertoire arising out of resolutionM50.11 above.</p>	
n.	<p><i>M50.14 (Qur'anic Arabic letters):</i> With reference to documents <a href="#">N3185</a> and <a href="#">N3215</a> item 4, WG2 accepts to encode the following combining characters:  0617 ARABIC SMALL HIGH ZAIN  0618 ARABIC SMALL FATHA  0619 ARABIC SMALL DAMMA  061A ARABIC SMALL KASRA  with their glyphs as shown in the code chart on page 7 in document <a href="#">N3211</a>.</p>	
o.	<p><i>M50.15 (Old Persian and Azerbaijani characters):</i> With reference to documents <a href="#">N3180</a> and <a href="#">N3215</a> item 4, WG2 accepts to encode 8 Arabic characters at code positions 0616, and 063B to 063F in the Arabic block, and 077E and 077F in the Arabic Supplement block, with their names and glyphs as shown on pages 7 to 10 in document <a href="#">N3211</a>. The first one of these is a combining mark.</p>	
p.	<p><i>M50.16 (Mathematical and Symbol characters):</i> With reference to documents <a href="#">N3198</a>, WG2 accepts to encode 29 mathematical and symbol characters at code positions:  2064 in the General Punctuation block, with its name and glyph from pages 22 and 23 in document <a href="#">N3211</a>  20F0 in the Combining Diacritical Marks for Symbols block, with its name and glyph from pages 24 and 25 in document <a href="#">N3211</a>  27CC, 27EE and 27EF in the Miscellaneous Mathematical Symbols block, with their names and glyphs from pages 30 and 31 in document <a href="#">N3211</a>  2B1B to 2B1F, 2B24 to 2B2F 2B45, 2B46, and 2B50 to 2B54 in the Miscellaneous Symbols and Arrows block, with their names and glyphs from pages 31 and 32 in document <a href="#">N3211</a>.</p>	
q.	<p><i>M50.17 (Symbol for Samaritan Source):</i> With reference to document <a href="#">N3217</a>, WG2 accepts to encode:  214F SYMBOL FOR SAMARITAN SOURCE  in the Letterlike Symbols block, with its glyph as shown on page 25 in document <a href="#">N3211</a>.</p>	
r.	<p><i>M50.18 (Latin capital letter sharp S):</i> With reference to document <a href="#">N3227R</a>, WG2 accepts to encode:  1E9E LATIN CAPITAL LETTER SHARP S  in the Latin Extended Additional block, with its glyph as shown in document <a href="#">N3227R</a>.</p>	
s.	<p><i>M50.19 (Bopomofo character):</i> With reference to documents <a href="#">N3179</a> and <a href="#">N3215</a> item 2, WG2 accepts to encode:  312D BOPOMOFO LETTER IH  with its glyph as shown on page 36 in document <a href="#">N3211</a>.</p>	

Item	Assigned to / action (Reference resolutions in document <a href="#">N3254</a> , and unconfirmed minutes in document <a href="#">N3253</a> for meeting 50 - this document you are reading)	Status
t.	<i>M50.20 (Disunification of CJK ideograph 4039):</i> With reference to documents <a href="#">N3196</a> , WG2 accepts to encode the new CJK Ideograph at code position 9FC3, with the second glyph in the figure on page 1 of document <a href="#">N3196</a> , and the source references G_KX0809.020, KP1-5E2B and T4-3946. WG2 also accepts to add a new source reference T6-4B7A to CJK ideograph at code position 4039 (keeping its current source references G3-5952 and KP1-5E34 and removing its source reference T4-3946).	
u.	<i>M50.21 (Orthographic and Modifier characters):</i> With reference to document <a href="#">N3216</a> , WG2 accepts to encode the following 4 combining marks: A789 MODIFIER LETTER COLON A78A MODIFIER LETTER SHORT EQUALS SIGN A78B LATIN CAPITAL LETTER SALTILLO A78C LATIN SMALL LETTER SALTILLO in the Latin Extended-D block, with their glyphs as shown on page 41 in document <a href="#">N3211</a> .	
v.	<i>M50.22 (3 Combining Macrons):</i> With reference to document <a href="#">N3222</a> , WG2 accepts to encode the following 3 combining marks: FE24 COMBINING MACRON LEFT HALF FE25 COMBINING MACRON RIGHT HALF FE26 COMBINING CONJOINING MACRON in the Combining Half Marks block, with their glyphs as shown in document <a href="#">N3222</a> .	
w.	<i>M50.23 (Indic characters for Vedic):</i> With reference to document <a href="#">N3235R</a> , WG2 accepts to encode the following 4 combining marks: 0B44 ORIYA VOWEL SIGN VOCALIC RR 0B62 ORIYA VOWEL SIGN VOCALIC L 0B63 ORIYA VOWEL SIGN VOCALIC LL in the Oriya block, and, 0D63 MALAYALAM VOWEL SIGN VOCALIC LL in the Malayalam block with their glyphs as shown in document <a href="#">N3264</a> .	
x.	<i>M50.24 (Ancient Roman characters):</i> With reference to documents <a href="#">N3218</a> and <a href="#">N3215</a> item 6, WG2 accepts to encode 11 additional characters at code positions: 2185 to 2188 in the Number Forms block, 2E19 in the Supplemental Punctuation block, A7FB to A7FF in the Latin Extended-D block, and 1019B in the Ancient Symbols block, with their names and glyphs as shown in document <a href="#">N3211</a> .	
y.	<i>M50.25 (CJK Unified Ideographs additions):</i> WG2 accepts to encode the 7 CJK Unicode-sourced ideographs, with their glyphs as shown in document <a href="#">N3210</a> at code positions 9FBC to 9FC2.	
z.	<i>M50.26 (Medievalist and Iranianist punctuations):</i> With reference to document <a href="#">N3193</a> , WG2 accepts to encode the following 20 characters at code positions 2E1A, 2E1B, 2E1E, 2E1F, 2E2A, 2E2C, 2E2E, 2E2F, 2E34, 2E38, and 2E40 to 2E49 in the Supplemental Punctuation block, with their names and glyphs as shown on pages 34 and 35 in document <a href="#">N3211</a> .	
aa.	<i>M50.27 (Progression of Amendment 4):</i> WG2 resolves to include all the items accepted for inclusion in the standard noted in resolutions M50.10 to M50.26 above into Amendment 4. WG2 instructs its project editor to forward the final text of Amendment 4 along with the disposition of comments document <a href="#">N3261</a> to the SC2 secretariat for an FPDAM ballot. The final set of charts and names lists are in document <a href="#">N3264</a> . The corrected starting dates for this work item are: FPDAM 2007-06-01, and FDAM 2007-12.	
bb.	<i>M50.28 (Archaic Sinhala numbers):</i> With reference to document <a href="#">N3195R</a> , WG2 accepts to encode 20 characters in the Sinhala block at code positions 0DE7 to 0DEF and 0DF5 to 0DFF, with their names and glyphs as shown on pages 3 and 4 in document <a href="#">N3195R</a> .	
cc.	<i>M50.29 (Egyptian Hieroglyphs):</i> With reference to document <a href="#">N3237</a> , WG2 accepts to encode in the standard 1063 hieroglyphs in code positions 13000 to 13426 in a new block 13000 to 1342F named Egyptian Hieroglyphs, with their names and glyphs as shown in document <a href="#">N3237</a> .	

Item	Assigned to / action (Reference resolutions in document <a href="#">N3254</a> , and unconfirmed minutes in document <a href="#">N3253</a> for meeting 50 - this document you are reading)	Status
dd.	<i>M50.30 (Coptic additions):</i> With reference to documents <a href="#">N3222</a> , WG2 accepts to encode in the standard 7 characters in code positions 2CEB to 2CF1 in the Coptic block, with their names and glyphs as shown in document <a href="#">N3222</a> . Three of these characters are combining marks.	
ee.	<i>M50.31 (Tai Viet script):</i> With reference to documents <a href="#">N3220</a> and <a href="#">N3221</a> , WG2 accepts to encode in the standard 73 characters in code positions AA80 to AAC3 and AADB to AADF in a new block AA80 to AADF named Tai Viet, with their names and glyphs as shown in document <a href="#">N3220</a> . Several of the characters are combining marks.	
ff.	<i>M50.32 (Bamum script):</i> With reference to document <a href="#">N3209</a> , WG2 accepts to encode in the standard 88 characters in code positions A6A0 to A6F7 in a new block A6A0 to A6FF named Bamum, with their names and glyphs as shown in document <a href="#">N3265</a> . Two of these characters are combining marks.	
gg.	<i>M50.33 (Meitei Mayek script):</i> WG2 accepts to encode in the standard 76 characters in code positions 1C80 to 1CBC, 1CBF to 1CCC and 1CCF, in a new block 1C80 to 1CCF named Meitei Mayek, with their names and glyphs as shown in document <a href="#">N3206</a> .	
hh.	<i>M50.34 (Hangul Jamo additions):</i> WG2 accepts to encode in the standard 117 Hangul Jamo characters as detailed in document <a href="#">N3242</a> . This includes creating two new blocks A960 to A97F named Hangul Jamo Extended-A and D7B0 to D7FF named Hangul Jamo Extended-B along with annotations proposed in document <a href="#">N3172</a> .	
ii.	<i>M50.35 (Avestan script):</i> With reference to document <a href="#">N3197</a> , WG2 accepts to encode in the standard 62 characters in code positions 10B00 to 10B35 and 10B38 to 10B3F in a new block 10B00 to 10B3F named Avestan, with their names and glyphs as shown in document <a href="#">N3197</a> .	
jj.	<i>M50.36 (Amendment 5 – subdivision and PDAM text):</i> WG2 instructs its editor to prepare a project sub division proposal and PDAM text based on resolutions M50.27 to M50.35 above, and forward them to the SC2 secretariat for ballot. The proposed start dates for the progression of this work item are: PDAM 2007-05-10, FPDAM 2007-10-30, and FDAM 2008-07. Due to the tight schedule prior to meeting M51 national bodies are encouraged to copy the project editor on their ballot comments at the same time they submit them to SC2 secretariat.	
<b>AI-50-4</b>	<b>Project Editor, Michel Suignard</b>	
	<i>To act on the resolution below.</i>	
a.	<i>M50.37 (WD of next edition):</i> With reference to documents <a href="#">N3229</a> , <a href="#">N3230</a> , <a href="#">N3248</a> and <a href="#">N3214</a> , WG2 instructs its editor to prepare working draft of the next edition of the standard for consideration at meeting M51, incorporating the texts of Amendments 1 through Amendment 4, along with a rationale document for the proposed changes in document <a href="#">N3229</a> , and including the expanded names list based on the <i>NamesList.txt</i> file publicly available from the Unicode Consortium's web site.	
<b>AI-50-5</b>	<b>Ad hoc group on principles and procedures (lead - Dr. V.S. UMaaheswaran)</b>	
	<i>To take note of and act upon the following items.</i>	
a.	Check and propose any updates to the P&P document arising from adopting the combined code table and enhanced nameslist format per document <a href="#">N3214</a> .	
<b>AI-50-6</b>	<b>Chinese national body - Mr. Chen Zhuang and Mr. Ngodrup</b>	
	<i>To take note of and act upon the following items.</i>	
a.	<i>M50.40 (Addition of 3 Tibetan characters):</i> WG2 invites the Chinese national body to discuss and resolve the concerns raised at meeting M50 on the proposal for 3 additional Tibetan characters in document <a href="#">N3240</a> by other WG2 experts.	
<b>AI-50-7</b>	<b>China, Ireland and UK - national bodies; Prof. Silamu, Mr. Michael Everson and Mr. Andrew West</b>	
	<i>To take note of and act upon the following items.</i>	
a.	<i>M50.39 (Orkhon):</i> With reference to document <a href="#">N3258</a> , WG2 invites the national bodies of China, Ireland and the UK to prepare a revised proposal taking into account the discussion at meeting M50 on Orkhon, for consideration at meeting M51.	
<b>AI-50-8</b>	All national bodies and liaison organizations	
	<i>To take note of and act upon the following items.</i>	

Item	Assigned to / action (Reference resolutions in document <a href="#">N3254</a> , and unconfirmed minutes in document <a href="#">N3253</a> for meeting 50 - this document you are reading)	Status
a.	<i>M50.38 (Hangul Jamo usage model)</i> : With reference to document <a href="#">N3257</a> , WG2 notes that national body of Republic of Korea has withdrawn their proposal for a new Hangul Jamo usage model in document <a href="#">N3113</a> .	
b.	<i>M50.39 (Orkhon)</i> : With reference to document <a href="#">N3258</a> , WG2 invites the national bodies of China, Ireland and the UK to prepare a revised proposal taking into account the discussion at meeting M50 on Orkhon, for consideration at meeting M51.	
c.	<i>M50.40 (Addition of 3 Tibetan characters)</i> : WG2 invites the Chinese national body to discuss and resolve the concerns raised at meeting M50 on the proposal for 3 additional Tibetan characters in document <a href="#">N3240</a> by other WG2 experts.	
d.	To review and feedback on the following contributions introduced at meeting M49: <a href="#">N3241</a> Parthian, Inscriptional Pahlavi, and Psalter Pahlavi scripts <a href="#">N3236</a> Anatolian Hieroglyphs <a href="#">N3235</a> Vedic Sanskrit	
e.	<i>M50.43 (Future meetings)</i> : WG 2 meetings: Meeting 51 - 2007-09-17/21, Hangzhou, China (note new location); along with SC2 plenary. Meeting 52 - 2008-05-19/23 (to be confirmed), West Coast, USA. Meeting 53 - 2008-10, (tentative) Tokushima (near Osaka), Japan; along with SC2 plenary. Meeting 54 - 2009-05, (tentative) Hong Kong SAR. Meeting 55 - 2009-10 - seeking host. IRG meetings: IRG #28 - 2007-06-04/08, Xi'an, China (note new location) IRG #29 - 2007-11-12/16, San Jose, California, USA (host Adobe; to be confirmed).	
f.	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. Manichaeian script – document <a href="#">N2544</a> Dictionary Symbols – document <a href="#">N2655</a> Samaritan Pointing characters – document <a href="#">N2758</a> Babylonian Pointing characters – document <a href="#">N2759</a> Bantu Phonetic Click characters – document <a href="#">N2790</a> Invisible Letter – document <a href="#">N2822</a> Palestinian Pointing characters - document <a href="#">N2838</a> Kaithi script -- document <a href="#">N3014</a>	

**End of Minutes**