

**Resolutions from the SC2/WG2 meeting in London, September 21-25, 1998**  
with comments from Ken Whistler, September 29, 1998

**Resolution M35.1 (FPDAM-18 on Symbols and Other characters including the EURO).**

This one survived basically intact. (No, the EURO did not get moved!!)  
The changes were:

237C SQUARE WAVE was removed (deemed to be a dup of 238D MONOSTABLE SYMBOL).  
This was responsive to the U.S., U.K., and Ireland comments.

The defect report N1875 reported on the out-of-place position for  
U+2047 REVERSED QUESTION MARK (which came from the ISO 2047:1975 pictures  
for control codes -- UCS code position and originating standard  
number were a coincidence, by the way). The character was moved  
to U+2426 and renamed SYMBOL FOR SUBSTITUTE FORM TWO.

In addition, as Asmus reported, there was a tussle over the "Romanian"  
characters: S/T with comma below. Asmus worked this out offline to  
produce clarifying text which can be added to Annex P in 10646.

FPDAM-18 progresses to DAM ballot.

=====

**Resolution M35.2 (FPDAM-21 on Sinhala)**

The Ireland NB comments to restore the four nuktated letters for  
expressing Tamil in the Sinhala script were accepted. This added:

U+0DB2 SINHALA LETTER NNNAYANNA  
U+0DBB SINHALA LETTER RRAYANNA  
U+0DBE SINHALA LETTER LLAYANNA  
U+0DBF SINHALA LETTER LLLAYANNA

A minor font fix was approved. Otherwise the encoding is unchanged.

FPDAM-21 progresses to DAM ballot.

=====

**Resolution M35.3 (PDAM-22 on Keyboard Symbols)**

The complex U.S. comments on changing names and changing the positions  
of other characters were accepted (cf. L2/98-064). (I won't bother  
recapitulating all of them here -- a revised listing will be available  
once the database is updated.) Several other name changes from Ireland  
and U.K. were accepted. And the German request to modify the glyph for  
what has become U+2396 DECIMAL SEPARATOR KEY SYMBOL was accomodated  
in part. The missing character U+239A CLEAR SCREEN SYMBOL was restored.

The upshot: this one got completely stirred by the ballot comments,  
so for sure do not implement based on the PDAM text!

PDAM-22 progresses to FPDAM ballot.

=====  
**Resolution M35.4 (PDAM-24 on Thaana)**

This one moved ahead unchanged, despite rumbles that one or more parties in the Maldives claim a different alphabetic order. In the absence of any concrete feedback, and with all positive votes, this one sails on.

PDAM-24 progresses to FPDAM ballot.

=====  
**Resolution M35.5 (PDAM-25 on Khmer)**

Japan's comments raised a number of issues on the PDAM. These were effectively countered by Maurice Bauhahn, who came to the meeting to explain and defend the Khmer proposal. The explanations satisfied Japan, which reversed its NO vote. There were a number of character name changes approved (Ireland comments, originating from Bauhahn, ultimately). The name of the currency was changed from RIAL to RIEL, but because it is based on the Khmer RA character rather than a Latin letter, it was not moved into the currency block. U+17C7 and U+17C8 will be changed to have combining circles shown and will be added to the list of combining characters (U.S. comment).

PDAM-25 progresses to FPDAM ballot.

=====  
**Resolution M35.6 (FPDAM on Syriac)**

The three combining marks for general Arabic use were moved and renamed as follows.

old:

U+074B MADDAH ABOVE  
U+074C HAMZA ABOVE  
U+074D HAMZA BELOW

new:

U+0653 ARABIC MADDAH ABOVE  
U+0654 ARABIC HAMZA ABOVE  
U+0655 ARABIC HAMZA BELOW

This satisfied U.S., U.K., and Ireland comments.

This change may have an impact on decompositions for Arabic characters in the Unicode Standard.

5 other character name misspellings were noted for correction.

The missing list of combining characters for 10646 Annex B will be added.

FPDAM-27 progresses to DAM ballot.

=====  
**Resolution M35.7 (FPDAM-17 on CJK Extension A)**

The discussion on this one centered mostly on niceties of layout for publication. There were several small additions to text for 10646 (Clause 26, Annex T, Annex L) approved or approved in principle. These have no impact on the Unicode Standard. The list of characters and their positions remains unchanged.

China is on record as follows regarding the fonts:

"The National Body of China has agreed to make arrangements for the provision of the necessary fonts."

FPDAM-17 progresses to DAM ballot.

=====  
**Resolution M35.8 (FPDAM-23 on Bopomofo Extended and other characters)**

The glyphs for 20E2 and 20E3 were interchanged in the FPDAM, and will be fixed (and enlarged a bit if possible).

The name of U+1673 was corrected to NNGO (from NGO).

The two new modifier letters at U+02EA and U+02EB had "MODIFIER LETTER" inserted at the beginning of their names for consistency with the other tone marks.

There were other minor editorial fixes to the text of the amendment.

FPDAM-23 progresses to DAM ballot.

=====  
**Resolution M35.9 (Procedures for character set registration)**

This one occasioned much discussion. It turns out the procedures document for working with ISO 2375 (the pertinent standard for registering character sets) was authored by Hekimi years ago, in the days when registrations were done with LetraSet. Everyone agreed that it is in desperate need of updating. It was less clear that everyone heard and understood the issue of needing careful review of mapping to 10646. At any rate, the U.S. recommendations regarding character set registration procedures and requesting resurrection of a Registration Advisory Group were endorsed and passed on to SC2.

=====  
**Resolution M35.10 (Defect report on Zones)**

Asmus already reported on this one. Rather than keep making minor (and pointless) editorial corrections to zone boundaries and zone diagrams in 10646, the zones will simply be removed in the next editorial corrections for the republication of 10646. The intended function of the zones has been overtaken by the much more detailed and explicit Roadmap now incorporated in the WG2 Principles and Procedures document.

=====  
**Resolution M35.11 (Mongolian script)**

Mongolian finally (!) progressed to ballot. The questions I raised at the Seattle meeting were ironed out to the satisfaction and mutual agreement of the Chinese delegation, the Mongolian delegation, and outside Mongolian experts attending the meeting.

The Mongolian character names, shapes and relative positions are basically unchanged from previous documents seen by UTC and L2. The punctuation/symbols area of the proposal got compressed a little, to account for some control characters removed from the proposal or encoded elsewhere.

The big change is that Mongolian is encoded at 1800..18AF, instead of starting at 1000. This is because it takes 11 columns, whereas only 10 columns were available starting at 1000. Rather than completely reorder the Mongolian proposal and remove all the holes in the punctuation/symbols area to squeeze it into 10 columns, it seemed better to leave some space for additional Mongolian symbols which might show up (cf. the Tibetan extensions), and code the block in a space which would minimize the rearranging required for the proposal.

The Mongolian space got moved to general punctuation and renamed:

202F NARROW NO BREAK SPACE

The question exclamation mark got moved to general punctuation, and its inverted form was cloned as well:

2048 QUESTION EXCLAMATION MARK

2049 EXCLAMATION QUESTION MARK

The four proposed positional format control characters were dropped in favor of use of ZWJ and ZWNJ, as for Arabic.

The three Mongolian variant selector characters were encoded, and stay in the Mongolian block.

The Mongolian vowel separator character is encoded, and the table of forms (documenting the use of the variant selector characters) will be adjusted to converge two forms made redundant by the presence of the vowel separator character.

The Mongolian proposal progresses directly to FPDAM ballot. (Amendment 29 ?)

=====  
**Resolution M35.12 (Additional Latin and other characters)**

This is "Bucket 35", the collection of all the odds and ends agreed to for encoding at WG2 meeting 35. Effectively it is just a list of letters and symbols grouped together, so that only one ballot is required, rather than several.

Altogether 66 new characters were dumped into this bucket. I won't list all the individual code points and character names here, but the various proposals that were dealt with included:

1. N1840 [L2/98-293] (The U.S. response to the TC-46 derived character proposals),  
N1885 [L2/98-292] (The U.S. response to several other Everson proposals),  
N1847 (Everson's response to N1840 and N1885)

The ad hoc bucket brigade met and picked out the uncontentious ones from this morass. 28 characters in toto, which largely reflected what the U.S. position paper had noted as o.k. to encode. 10 characters were completely rejected as unsuitable for encoding, and everything else in the entire suite of proposals was tossed into the category of "requires more study". The 28 characters which made it in include:

Latin letter ou  
Latin letter z with hook  
Greek kai symbol  
Greek small letters stigma, digamma, koppa, sampi  
Cyrillic hundred thousands sign and millions sign  
10 Cyrillic characters for Kildin Sami  
the Tironian sign et  
Roman numeral reversed one hundred  
5 other symbols/punctuation signs

2. N1332 10 precomposed Latin letters for Livonian

Note: Some of these have two diacritics, but they do not introduce any problems for normalization.

3. N1817 Modifier letter double apostrophe (for Nenets)
4. N1812 Modifier letter cross accent (for Swedish)
5. N1838 4 binary completion (Latin) letters (for normalization)

By the way, Klaas Ruppel affirmed that a with dot above absolutely is used in Uralic orthographies for Sami, so this is not a bogus letter.

6. N1857 Mongolian currency sign:

20AE TUGRIK SIGN

7. N1845 IPA characters for disturbed speech

This set was winnowed down to 5 base forms, 2 modifier letters, and 10 combining marks.

8. N1882 Interlinear annotation characters:

FFF9 INTERLINEAR ANNOTATION ANCHOR  
FFFA INTERLINEAR ANNOTATION SEPARATOR  
FFFB INTERLINEAR ANNOTATION TERMINATOR

These are the "Ruby" annotation characters. (Cf. L2/98-099, etc.)  
The Japanese delegation raised questions about these in N1861, pointing out that they were insufficient for a general representation of an "interlinear object" as defined by SC18. WG2 went with the more narrow interpretation, just following existing practice in Japanese ruby implementations.

9. N1886 conditional space:

204F SOFT SPACE

This is a new critter introduced at the behest of Khmer. It would be used for languages which prohibit word spacing but which include justification (Khmer, Myanmar, Lao, Thai, etc.).

Bucket 35 progresses directly to FPDAM ballot (Amendment 30?)

=====

**Resolution M35.13 (PDAM-14 on Yi script)**

The problems with the Yi encoding also required an ad hoc to sort through and resolve. The problem is basically that the Yi proposal originating with China followed the order of China's national standard for Yi. The agreement in WG2 had been to follow the traditional Yi order as documented by Bburx. The new Chinese proposal which came in accompanying their NO vote on Yi fixed all the errors in the Yi ordering that the U.S. ballot comments noted (various stirrings of characters and an off-by-one error), but left the consonants in the Chinese standard order rather than the Yi order. The ad hoc agreed to:

1. Reorder the Yi chart one more time into the traditional Yi order as documented by Bburx. (China will amend its national standard to this order, so that it is in accord with 10646 when the dust settles.)
2. Agreed \*not\* to add any more Yi radicals.
3. Acquiesced in some names changes proposed by China, and questioned four more names that China could not verify at the meeting.

Amendment 14 on Yi progresses to FPDAM ballot.

=====  
**Resolution M35.14 (PDAM-26 on Myanmar/Burmese)**

Myanmar got the most attention of any script at WG2. It took 4 days of ad hoc meetings, off and on. Myanmar sent a 5-member delegation--all of whom turned out to be highly qualified computer professionals. The Myanmar delegation brought in a proposal completely different from the PDAM text--a Myanmar encoding they call "autolexicography" that serves for Myanmar input and rendering, but which also has an order and dummy characters encoded so as to result in auto-sorting of data as well.

After much discussion, the issue of sorting Myanmar was separated from encoding in 10646. As it turns out, the Myanmar proposal was assuming the virama model for handling subjoined consonants (they were using a "LINK" character for these), so the ad hoc fairly quickly was able to converge on the basic encoding. The problem areas turned out to be just how many vowels and dependent vowel signs to encode. Eventually everybody came to an acceptable compromise for the FPDAM.

The Japanese ballot asked for separate encoding of all the subjoined forms (essentially rejecting the virama model). In the face of the Myanmar delegation's expressed preferences, however, and all the other expert opinion, they acquiesced to the ad hoc consensus and agreed to change their vote to YES.

The significant points of change from the PDAM text are:

1. 4 complex vowels and vowel signs (aa, ai, -o, -au) were removed as redundant. However, space was left in the chart for them so that if implementation experience proves it expedient to encode them as units, their natural positions will be available. Space was also left for ui -ui as well, although they were not encoded.
2. The kinzi character was removed. It is a conjunct form.
3. The virama and killer were unified to follow the more generic Indic model for explicit virama. This will require use of the ZWNJ to express the visible killer, but was deemed more acceptable to Myanmar and some others present. It is less confusing than the two forms present in the PDAM. (Should implementation experience prove this infeasible, the explicit killer character can be added back later.)
4. The Myanmar character AFOREMENTIONED (U+174F in the PDAM text) was modified from the glyph fragment shown there to the full, normal form of the particle.
5. The extension characters for Pali/Sanskrit were left in, despite the allegation from Myanmar that those are "never" seen. The desire to separate them off from Basic Myanmar characters was accommodated by defining two distinct collections for Myanmar (but only a single block for the entire script).
6. A new font will be used for the FPDAM that follows the changes outlined by Zaw Htut in his paper and supported by Myanmar.

7. Myanmar indicated a strong desire to change the script and character names from "Burmese" to "Myanmar", so the block name and all character names will be updated.
8. To accomodate the probable number of Myanmar extension characters required in the future to cover Kayin, Mon, Shan, etc., the entire Myanmar block was moved from U+1700..U+177F to U+1000..U+109F (the area vacated by Mongolian). That allowed an additional 2 columns, and makes it likely that all the minority language additions can be handled in a contiguous chunk.

The longer discussion of all this is available in WG2 N 1884R2.

Amendment 26 for Myanmar (Burmese) progresses to FPDAM ballot.

=====

**Resolution M35.15 (IRG - Extension B Format)**

This resolution has no impact on what characters get encoded, but put WG2 on record as preferring a single column format for the CJK extensions in the works for Plane 2. This enormously simplifies the printing task for the IRG and makes it likely that Extension B will be concluded more quickly.

=====

**Resolution M35.16 (PDAM-28 - Ideographic Description Characters)**

This is the set of 12 ideographic description symbols that have been kicked around for several years now. The entire set was encoded. They were encoded at U+2FF0..U+2FFF, which was a "dead air" space at the end of the Kangxi radicals, just before the CJK punctuation.

The names of 8 of these characters were changed to make them more comprehensible in English.

Their function is as explained at the Seattle meeting. They are used with CJK characters from the URO (or the new set of radicals) to describe Han characters. But they do not create equivalences with existing encoded Han characters.

Amendment 28 for Ideographic Description Characters progresses to PDAM ballot.

=====

**Resolution M35.17 (PDAM-15 - KangXi [and other CJK] radicals)**

This PDAM has been delayed by the uncertain status of the set of additional CJK radicals included. That started at 32 and eventually ballooned to the current list of 116.

The 214 + 2 KangXi radicals are encoded at U+2F00..U+2FDF.  
The 116 additional CJK radicals were original proposed for U+2E00..U+2E7F, but were moved to U+2E80..U+2EFF, so that they are contiguous with the KangXi radicals and so that no one gets the impression that the intervening 8 columns could be used for yet more CJK radicals that somebody might discover.

Amendment 15 for KangXi and other CJK radicals progressed to PDAM ballot.

=====  
**Resolution M35.18 (PDAM on Tibetan Extension)**

The proposed text for Tibetan extensions was accepted virtually unchanged. Unlike other extensions, this amendment will be treated effectively as a complete replacement for Amendment 7 (Tibetan). This allows various defect reports on glyphs, combining status, etc., to be incorporated into the new ballot, rather than having to be processed separately. 25 new characters are added for Tibetan.

The Chinese delegation questioned several of the glyph fixes (most of which involved a reinterpretation of combining status), but because they had no Tibetan expert present, they could not take any decisions on the issues at the meeting. Chris Fynn, one of the Tibetan experts who participated extensively in the development of the Tibetan extension proposal, attended the meeting and wrote up a detailed response to China's concerns, complete with excerpts directly from Tibetan manuscripts he had brought to the meeting. But because China could not reach consensus, this one goes to PDAM instead of FPDAM.

The Tibetan Extensions proposal progresses to PDAM ballot (Amendment 31?)

=====  
**Resolution M35.20 (Future Meetings)**

The important thing to note here is that WG2 officially moved its next meeting to March 9-15 in Fukuoka, Japan. This was explicitly to remove the clash in times against the Unicode Conference scheduled in Boston, so that people who feel they must attend both can manage it.

=====  
end of document