

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 **N 1034**

Date: 1994-04-22

Source: WG 2 Meeting 25, Antalya, Turkey, 1994-04-17--22 Title: Resolutions of WG 2 meeting 25 Action: To be considered by SC 2 and For information to WG 2 Status: Adopted resolutions of WG 2 Distribution: ISO/IEC JTC 1/SC 2 and SC 2/WG 2

RESOLUTION pDAM-UTF-16:

UNANIMOUS

WG 2 instructs its editor to revise and update pDAM 1 in N 970R as follows:

- a. Editorial suggestions and Annex M of document N 976 from UK with the following further modifications to the definition of RC element:
 - i. 4.36 RC-element: a two-octet sequence comprising the R-octet and the C-octet (see 6.2) from the four-octet sequence that corresponds to a cell in the coding space of this coded character set.

Clause 14.1 "Two-octet BMP form":
Add at the end of paragraph 2:
"(i.e. its RC-element)".
 - ii. Paragraphs related to error handling as documented in N 1027 - 'UTF-16 Specification - X.6 Incorrect Sequences of S-Zone RC Elements'
- b. Comments expressed in document N 991
- c. Editorial clarification in Section 9 of the standard as drafted in document N 1021, addressing the concerns expressed in N 972 from Korea.

Further the editor is requested to forward the updated pDAM 1 to SC 2 secretariat for further processing, (by 1994-06-01 if possible, to enable SC 2 ballot results to be available to WG 2 by its meeting on 1994-10-10).

RESOLUTION pDAM-UTF-8:

UNANIMOUS with Amendments

WG 2 instructs its editor to revise, update and forward to SC 2 secretariat for further processing pDAM 2 in N 993 as follows:

- i. **Amendments to Draft Annex N, as in N 993**
 - a. **Move three paragraphs, after the title, to a Note at the end of the Annex.**
 - b. **In N.2 replace first sentence by :**

In UTF-8 the characters from this standard shall have coded representations that comprise sequences of octets of lengths 1, 2, 3, 4, 5 and 6 octets.
 - c. **In para 2, 3rd line of the table, add after 0000 FFFF:**

(excluding 0000 D800 to 0000 DFFF)
 - d. **In para 3, first item, add:**

Coded representations in the range 0000 D800 to 0000 DFFF shall be excluded from this conversion process.
 - e. **Para 2, line 1, add "coded" before "character"**
Para 3, line 3, delete "character".
Para 4, item 1, line 3, replace "character code" by "coded representation"
 - f. **Amend all text of N.2 to Normative style, using "shall".**
 - g. **In N.1 add new bullet :**
 - **The transformation defined for UTF-8 is applicable to characters coded according to UCS-4 or UCS-2, but not for characters coded according to UTF-16.**

If it is desired to transform data coded in UTF-16, this can be accomplished first by transforming from UTF-16 to UCS-4 (see ____) before applying UTF-8.
- ii. **Remove all references to UTF-1 from pDAM 2 subject to the acceptance of Resolution UTF-1.**
- iii. **Add the following text to clarify handling of incorrect octet sequences in the UTF-8 form:**

nn Incorrect Sequences of octets:
Interpretation by receiving devices

If a CC-data-element includes either:

 - a sequence of a first octet that is not immediately followed by correct number of following octets, or
 - one or more adjacent following octets that are not required to complete a sequence of first and following octets,

then according to Y.nn (nn) such a sequence of octets is not in conformance with the requirements of UTF-8. It is known as a malformed sequence, and it shall not be transmitted by an originating device that is in conformance with the requirements of UTF-8.

If a receiving device receives a malformed sequence, because of error conditions either:

 - in the originating device, or

- in the interchange between the originating and the receiving device,
or
 - in the receiving device itself,
- then it shall interpret that malformed sequence in the same way that it interprets a character that is outside the adopted subset that has been identified for the device (see 2.3c).

RESOLUTION UTF-1:
WG 2 instructs its editor to

APPROVED - CANADA Abstains

- a. Prepare a new corrigendum to remove the informative Annex G on UCS transformation format (UTF-1)
- b. Transfer the current content of Annex G to ISO Registry following ISO 2375.

RESOLUTION PRIN-1: **UNANIMOUS**
WG 2 resolves that it will not add new characters to the code positions of ISO/IEC 10646-1, until pDAM 1 is formally approved by SC 2 letter ballot, except for the use as defined in pDAM 1.

RESOLUTION PRIN-2: **UNANIMOUS**
WG 2 accepts the proposed categorization of characters as proposed in document N 1002 and instructs its editor to revise and update N 946 accordingly.

RESOLUTION PRIN-3: **UNANIMOUS**
WG 2 reconfirms the guidelines, principles and procedures in document N 946 for allocation of new characters and scripts to ISO/IEC 10646-1. All requesters for additional characters in the standard should take note.

RESOLUTION DR-1: **UNANIMOUS**
WG 2 instructs its editor to prepare editorial corrigenda forms from the instructions to the editor contained in N 1031R. Further the editor is instructed to create the appropriate response documents and forward them together with the Corrigenda Forms to the SC 2 secretariat for forwarding to ITTF.

RESOLUTION DR-2: **UNANIMOUS**
WG 2 instructs its editor to revise and update the Corrigendum 1 in N 994 as follows:

- a. Change the words 'editorial correction' to 'correction' in the corrigendum text.
- b. Add a statement justifying the corrigendum prior to the corrigendum text, referencing and attaching the defect reports in N 910, 911, 912 and 923 to the text.

Further the editor is instructed to forward it to SC 2 secretariat for further processing as a Technical Corrigendum, along with the response forms for the defect report.

RESOLUTION pDAM3: **UNANIMOUS**
WG 2 instructs its editor to prepare text for pDAM3 based on the text in N 995, and forward it to SC 2 secretariat for further processing as a pDAM.

RESOLUTION ADMIN-1: **UNANIMOUS**
WG 2 expresses its sincere thanks to Mr. Shigenobu Kato for his role as the rapporteur of the Ideographic Rapporteur Group (IRG) and wishes him all success.

RESOLUTION ADMIN-2: **UNANIMOUS**
WG 2 endorses the nomination of Mr. Akio Kido as the new rapporteur of IRG, and wishes him all success.

RESOLUTION ADMIN-3: **UNANIMOUS**
WG 2 expresses its sincere thanks to Mr. Ahmet Kurtur and all other members of the staff of Turkish Standards Institution(TSE) for the excellent arrangements, their kind hospitality and continued support during the WG 2 meeting.

RESOLUTION ADMIN-4: **UNANIMOUS**
WG 2 thanks Mr. Joachim Friemelt for his valuable contributions to the work of WG 2 over the past several years and wishes him a happy retirement.

RESOLUTION ADMIN-5: **UNANIMOUS**
WG 2 thanks Dr. V.S. UMaaheswaran for serving as meeting secretary for the Antalya WG 2 meeting.

RESOLUTION ADMIN-6: **UNANIMOUS**
WG 2 thanks the following for their informative presentations on different scripts at the meeting:

- i. Mr. N. Subramanian, CDAC, Pune, India, on Indian Standard Code for Information Interchange, ISCII, and its handling of the different scripts used in India
- ii. Professor Enghebatu, Mongolian Language Institute, China, on Mongolian script
- iii. Professor Nyima Trashi, Tibet University, China, on Tibetan script
- iv. Professor Wushur Slamu, Xianjiang University, China, on Uighur, Kazhak and Kirghiz scripts, and,
- v. Mr. YongJun Zhou, China Great-Wall Computer Group and Mr. Shamalayi, National Institute of Chengdu, China, on Yi script.

FUTURE MEETINGS:

- a. WG 2 meeting no. 26:
Week of 1994-10-10 (October 10, 1994) (Monday to Friday 5 days) in San Francisco Bay Area, USA; Hosted by The Unicode Consortium.
(Note: WG 2/IRG is planning to meet during the week of 1994-08-29 in Beijing, and SC 22/ WG 20 is meeting the week of 1994-10-03 in the US West Coast.)
- b. WG 2 meeting no. 27:
Week of 1995-06-26 (June 26, 1995) (Monday to Wednesday, 3 days), along with SC 2 plenary and WG 3 meetings in Helsinki, Finland.
- c. WG 2 meeting no. 28:
Week of 1995-11-06 (November 6, 1995) in Japan.
- d. WG 2 meeting no. 29:
Week of 1996-04-xx (April , 1996) (Monday to Friday, 5 days)
Canada or Denmark

===== End of Resolutions =====