Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

ISO/IEC JTC1/SC2/WG2
Universal Multiple-Octet Coded Character Set (UCS) – ISO/IEC 10646
Secretariat: ANSI

Doc Type: **Expert Contribution**
Title: **Revisions to Annex N of ISO/IEC 10646-1**
Source: **Michael Everson, EGT (IE)**
Project: **JTC1.02.18.01**
Status: **Proposal**
Date: **1998-10-17**
Distribution: WG2
Medium: **PDF**

Because of the extremely successful work that WG2 and the UTC have done in the years since the publication of the UCS, annex N of ISO/IEC 10646 is now out of date. Of the thirteen entries published in 1993, only two of them remain:

- ~~Burmese~~
- ~~Cree and Inuktitut~~
- ~~Ethiopian~~
- Extensions to various scripts for Indo-European languages
- Hieroglyphics
- ~~Khmer~~
- ~~Maldivian~~
- ~~Mongolian~~
- ~~Runic~~
- ~~Sinhalese~~
- ~~Syriac~~
- ~~Tibetan~~
- ~~Yi~~

It is suggested that WG2 delete the first of these as the original proposal (Nxxx) which led to its inclusion has not been followed up upon by its proponents, and in the estimation of numerous experts in the UTC and WG2, it is inadequately supported by a strong proposal for the addition of these characters. This would leave annex N with but a single entry, for a script which has (N1636 and N1637) already been proposed for inclusion in the UCS.

Annex N still has, however, an important use. In the Roadmaps for both Plane 0 and Plane 1, a large number of scripts has been identified as possible candidates for inclusion in the UCS. The draft below includes, exhaustively, the scripts listed in the Roadmaps.

It is true that this listing in the Roadmap does not guarantee the acceptance of any script in the UCS – indeed, the proposed deletion of the Indo-European characters proposed above is indicative of the informative nature of annex N. It may be that WG2 will choose not to include a number of the scripts listed below in the new annex N of the second edition of ISO/IEC 10646-1.

What is presented, however, is a list of script names together with a brief parenthetical note intended to explain the nature of each entry. Entries in **bold face** have already been proposed (formally or in an exploratory document) to the UTC or to WG2 (so maybe they will already be in by the time we republish). Entries **<u>bold and underscored</u>** means the script has already been provisionally accepted for Part 2 (which means they probably should not go into annex N.

Members of WG2 are asked to consider this list, to propose deletions or additions to it, and to accept it and its format to replace the current annex N in the second edition of ISO/IEC 10646-1 and/or the first edition of ISO/IEC 10646-2.

(Proposed amendment of) Annex N
(informative)

Scripts under consideration for future editions of ISO/IEC 10646

In order to make sure that ISO/IEC 10646 is useful for people using scripts of the world, characters included in ISO/IEC 10646 were selected with input and feedback from national standards organization and/or qualified experts.

Some scripts and symbols were not included in this edition because sufficient input and feedback have not been provided during the preparation and review stages.

It is foreseen that character code positions for these scripts and symbols may be allocated when sufficient input and review is provided. Such scripts and symbols include:

Scripts possibly to be encoded in the BMP:

Aramaic *(extinct script used in the Middle East)*
Avestan *(extinct script used in Zoroastrian Persia)*
Batak *(script used in South Asia)*
Buginese *(script used in South Asia)*
Buhid *(script used in the Philippines)*
Cham *(script used in Viet Nam)*
Do Dai *(script used in China)*
Glagolitic *(ecclesiastical script used for Old Church Slavonic)*
Javanese *(script used in South Asia)*
Kayah Li *(script used in South Asia)*
Kirat/Limbu *(script used in South Asia)*
Laai Dai *(script used in China)*

Meitei/Manipuri *(script used in South Asia)*
Newari *(script used in Nepal)*
Old Hungarian *(extinct medieval script of Hungary)*
Old Xishuang Banna Dai *(script used in China)*
Pahawh Hmong *(script used by the Hmong people)*
Pahlavi *(extinct script used ancient Persia)*
'Phags-pa *(extinct script used in Central Asia)*
Phoenician *(extinct script used in Phoenicia)*
Samaritan *(extinct script used in the Middle East)*
Siddham *(ecclesiastical script used in Japan)*
South Arabian *(extinct script used in the Middle East)*
Tagalog *(extinct script used in the Philippines)*
Tagbanwa *(script used in the Philippines)*
Tifinagh *(script used by the Berber people)*

Scripts possibly to be encoded in the Plane 1:

Aegean scripts *(extinct scripts used in Cyprus and Crete)*
Aiha *(modern script used by a people who might be going to have lived a very long time from now in Northern California)*
Ahom *(extinct script used in South Asia)*
Aymara Ideographs *(extinct script used in South America)*
Aztec Pictograms *(extinct script used in North Amarica)*
Balinese *(script used in Bali, similar to*

Javanese)
Balti *(extinct script used in South Asia)*
Bamum *(script used in Africa)*
Blissymbols *(modern script used by people with severe communicative disabilities)*
Brahmi *(extinct script used in Northern India, precursor of all scripts of India and most of Southeast Asia)*
Buthakukye *(extinct script used in Albania)*
Byzantine Musical Symbols
Carian *(extinct script used in the Western Asia)*
Chakma *(extinct script used in South Asia)*
Chalukya *(extinct script used in South Asia)*
Chinook
Chola
Cirth *(modern script used by Dwarves)*
Deseret *(modern script invented for the English language)*
Elbasan *(extinct script used in Albania)*
Etruscan *(extinct script used in Etruria)*
Gothic *(extinct script used for the Gothic language)*
Granta *(extinct script used in Southern India; possibly unifiable with Tamil)*
Hanunóo *(script used in the Philippines)*
Hittite Cuneiform *(extinct script used in Anatolia)*
Indus Valley *(extinct script used in Northern India; undeciphered)*
Iberian *(extinct script used in Iberia)*
Jurchin *(extinct script used in China)*
Kaithi *(script used in South East Asia)*
Kauder/Micmac *(extinct script used in North America)*
Khamti *(extinct script used in South East Asia)*
Kharoshthi *(extinct script used in South and Central Asia)*
Khotanese
Kitan Large Script *(extinct script used in China)*
Kitan Small Script *(extinct script used in China)*
Kök Turki *(extinct script used in Central Asia, precursor of Old Hungarian)*
Landa *(extinct script used in South Asia)*
Lycian *(extinct script used in Western Asia)*
Lydian *(extinct script used in Western Asia)*
Maghrebi *(variant of the Arabic script)*
Mandaean *(ecclesiastical script used in the Middle East)*
Manichaean *(extinct script used in Central Asia)*
Mayan Hieroglyphs *(extinct script used in North America)*
Meroitic *(extinct script used in the Middle East)*
Modi *(script used in South Asia)*
Nabataean *(extinct script used in the Middle East)*
Naxi Ideographs *(script used in China)*
Nepali *(script used in Nepal; possibly unifiable with Devanagari)*

Numidian *(extinct script used in the Middle East, precursor of Tifinagh and possibly unifiable with it)*
Old Permic *(extinct script used for the Komi language)*
Old Persian Cuneiform *(extinct script used for Old Persian)*
Osmanya *(modern but extinct script invented for the Somali language)*
Palmyrene *(extinct script used in Western Asia)*
Parthian
Paucartambo *(extinct script used in South America)*
Phaistos Disk *(extinct script found on one document on Crete; undeciphered)*
Pollard Phonetic *(modern script used in China)*
Proto-Byblic *(extinct script used in the Middle East)*
Proto-Elamic *(extinct script used in the Middle East)*
Pyu/Tircul *(extinct script used in South East Asia)*
Rejang *(script used in South East Asia)*
Rongorongo *(extinct script used in Easter Island; undeciphered)*
Satavahana *(ecclesiastical script used in Northern India and Tibet)*
Shavian *(modern script invented for the English language)*
Sinaitic *(extinct script used in Western Asia)*
Sogdian *(extinct script used in Central Asia)*
Sumerian Pictograms *(extinct script used in Babylonia and Assyria, precursor to cuneiform)*
Sumero-Akkadian Cuneiform *(extinct script used in Babylonia and Assyria)*
Takri *(script used in South Asia)*
Tengwar *(modern script used by Elves)*
tlhIngan Hol pIqaD *(modern script used by Klingons)*
Tungut Ideographs *(extinct script used in China)*
Ugaritic Cuneiform *(extinct script used in the Middle East)*
Uighur *(extinct script used in Central Asia)*
Vai *(script used in Africa)*
Veso Bei
Western Musical Symbols
Woleai *(extinct script used in South East Asia)*

This list is not exhaustive. Other scripts and symbols as well as additional characters for the included scripts are expected to be included in future editions of ISO/IEC 10646.