

X3L2/96-102

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 1437

1996-08-06

Title: Report of 3rd International Mongolian Encoding Meeting  
Source: China, Mongolia, UNU/IIST  
Status: Meeting Report  
Action: For the consideration of SC 2/WG 2  
Distribution: ISO/IEC JTC 1/SC 2/WG 2, Liaison Organisations

### History

After the presentation of draft proposal N1368 (Title: Joint proposal draft on encoding Mongolian Character Set) at the WG2 Copenhagen meeting in April 21st-27th, 1996, an ad-hoc group was formed with the objective of analysing those aspects of the draft proposal which were not coherent with the general principles and rules of the ISO 10646 Universal Multiple-Octet Coded Character Set encoding.

The ad-hoc group concluded that the draft should be revised and recirculated. For this purpose, the ad-hoc group suggested an experts' meeting to be held in China in early August 1996. This proposal (document N1383) was unanimously accepted in the resolution of the Copenhagen meeting.

The Chinese Electronics Standardization Institute of the Chinese Ministry of Electronics Industry hosted the 3rd International Mongolian Encoding Meeting in Beijing from August 1st to August 6th, 1996.

### Document Status and Present Situation

Before the August meeting, a draft proposal brought forth by Mongolia (MNISM, Mongolian National University, UNU/IIST) was circulated in May/June 1996. This draft proposal contains a character table (*Mongolian Basic Character Set* [Table 1a]), a list of names (*Names of Mongolian Basic Character Set* [Table 1b]) and separate character lists for all included alphabets. This draft addressed one of the major points of concern of the N1368 proposal, its huge number of glyphs which was due to the inclusion of presentation forms.

Another document was made available by China at the beginning of the August meeting. This document contains an introductory text (*Proposal on Encoding Mongolian Script*) a character list (*Mongolian Canonical Characters*), a list of character names (*Names of Mongolian Canonical Characters*), a concordance of unified characters (*Unification of Mongolian canonical characters*) a concordance of basic forms and presentation forms (*Rules of the Mongolian Todo Xibe Manchu Presentations to singly present*) as well as documentation on the use of control characters (*The Usage of the Peculiar Symbols and the Control Symbols of Mongolian, Todo, Xibe, Manchu Script*) for producing glyph selections which cannot be predicted by the general rules governing the Mongolian writing system.

All participants (see the appendix for a detailed list) of the meeting expressed their continuous commitment and their willingness to contribute to this issue of historical importance.

## Discussion Items

During the review of the above-mentioned documents it was commonly acknowledged that the work done so far is a significant and substantial improvement over the joint proposal N1368. However, all participants concluded that the documents needed further improvement before they can be circulated. Nearly all improvements discussed (as well as minor technical corrections, like spelling errors etc.) could be agreed upon during this meeting; only the question of the typeface needs further research and discussion. The commonly agreed items are listed separately. All character numbers given here refer to the Mongolian Basic Character Set as agreed upon during the meeting.

Unanimously accepted items include:

- The name of the basic character table, previously *Mongolian Canonical Characters*, is changed to *Mongolian Basic Character Set*.
- Due to specific properties of the Mongolian script, three special control characters are introduced since the generic joiner and non-joiner symbols do not offer sufficient functionality for the selection of appropriate glyphs at a given position in a word.
- These three control codes are named MONGOLIAN FIRST IDENTIFIER (#172), MONGOLIAN SECOND IDENTIFIER (#173) and MONGOLIAN POSITIONAL SELECTOR (#174).
- The separated tables *Unification of Mongolian Canonical Characters* and *Rules of the Mongolian Todo Xibe Manchu Presentations* to singly present will be recompiled into one single table. A joint draft for such a table, named *Mongolian Reference Table*, was agreed upon during the meeting. It will be completed and made available for review before the end of August, 1996.
- Some character names are changed as follows:

No.	Name (old)	Name (new)
#005	MONGOLIAN HORIZONTAL COLON	MONGOLIAN COLON
#008	MONGOLIAN TODO VERTICAL SOFT HYPHEN	MONGOLIAN TODO SOFT HYPHEN
#052	MONGOLIAN LETTER CA	MONGOLIAN LETTER CHA.
#059	MONGOLIAN LETTER GAA	MONGOLIAN LETTER KHA.
#082	MONGOLIAN LETTER TODO CA	MONGOLIAN LETTER TODO CHA.
#170	MONGOLIAN LETTER MANCHU AG BINDU	MONGOLIAN LETTER MANCHU AG BHA
#172	MONGOLIAN LETTER FIRST IDENTIFICATION CONTROL SYMBOL	MONGOLIAN FIRST IDENTIFIER
#173	MONGOLIAN LETTER SECOND IDENTIFICATION CONTROL SYMBOL	MONGOLIAN SECOND IDENTIFIER
#174	MONGOLIAN LETTER POSITIONAL CONTROL SYMBOL	MONGOLIAN POSITIONAL SELECTOR
- A symbol to be added is the combined question mark exclamation mark [?!] which cannot be replaced by a sequence of the separate characters of the same name since they would not appear in this form in the Mongolian vertical line arrangement. This combinatory symbol (*niilmel temdeg* in Mongolian) is placed at position #007 and is named MONGOLIAN COMBINATORY SYMBOL. Its inverted form is included in the table of presentation forms.
- One Todo letter is added, its position is #085, its name is MONGOLIAN LETTER TODO YA.
- One character which can be theoretically decomposed into other characters but which is traditionally regarded as a character unit in Mongolian is to be included in

the table. The character is  $\mathcal{L}$ , its name is to be MONGOLIAN LETTER LHA, its position is #064.

- Four presentation forms of the birga (MONGOLIAN BIRGA, #002) are included in the table of presentation forms.
- The necessity of a glyph shape table associating elementary glyph forms (e.g. *shilbe*, *shud*, *gedes* etc.) with certain canonical letters (e.g. I, A, O etc.) was again acknowledged but due to limited time it could not be discussed in detail during this meeting. It will be discussed soon.
- The wording of the *Proposal on encoding Mongolian script* has been corrected or slightly modified in a number of places in order to achieve conformity with other adopted naming conventions for table names, etc.
- The order of control codes, punctuation marks, letters of various kinds and digits is rearranged as follows:

000---012	MONGOLIAN SPACE (other punctuation marks as before)
007	Combinatory symbol
016---025	digits
032---155	MONGOLIAN LETTER A, etc.
172	MONGOLIAN FIRST IDENTIFIER
173	MONGOLIAN SECOND IDENTIFIER
174	MONGOLIAN POSITIONAL SELECTOR

- The name "Xibe" is changed into "Sibe" in all instances.
- The complete documentation for the Mongolian encoding proposal comprises the following five texts and tables:

1. Mongolian Basic Character Set
2. Mongolian Basic Character Name List
3. The Usage of Identifier and Selector Symbols in Mongolian Writing

Supplement:

- 3.1. Mongolian Coding Conventions
- 3.2. Mongolian Reference Table

- Spelling errors (e.g. the occasional \*MANGHU instead of MANCHU) are corrected.

The above-mentioned items were reconfirmed after thorough review and the meeting was concluded in the general consensus that important progress has been made.

All participants appreciate the efforts made by the host for the accommodation and wish to thank the Chinese Electronics Standardization Institute for their efforts.

Beijing, August 6th, 1996

Bosor  
Chinese Delegation

O. Chilkhaasuren  
Mongolian Delegation

O. Corff  
UNU/IIST Delegation