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

Doc Type: **Working Group Document**
Title: **Preliminary proposal for encoding the Kayah Li script in the BMP of the UCS**
Source: **Michael Everson**
Status: **Individual Contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2006-01-11**

This is a preliminary proposal to encode the Kayah Li script in the BMP of the UCS.

The Kayah Li script is used to write the Kayah or Kayah Li script, which is member of Karen branch of the Sino-Tibetan language family. It is also known as Red Karen and Karenni. It is spoken by about 210,000 people in the Kayah and Karen states of Myanmar, and by some people in Thailand.

Structure

Kayah Li's Brahmic origins are known only from its ordering. It has been reanalyzed as an alphabet; most of the vowels are written with spacing letters, although some of them are written with diacritics applied to the letter ꤰ A. Consonants have no inherent /a/ vowel sound. Digits exist.

Ordering

An ordering based on the traditional order of the Brahmic script, but with differences from it, is used. Further study needs to be done to compare it with neighbouring scripts.

A. Administrative

1. Title
Proposal for encoding the Kayah Li script in the UCS.
2. Requester's name
Michael Everson
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2006-01-11
5. Requester's reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
No.
6b. More information will be provided later
Yes.

B. Technical – General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
Yes.

Proposed name of script
Kayah Li.
1b. The proposal is for addition of character(s) to an existing block
No.
1c. Name of the existing block
2. Number of characters in proposal
46
3. Proposed category (see section II, Character Categories)
Category B.1.
4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)
Level 3
4b. Is a rationale provided for the choice?
Yes.
4c. If YES, reference
Combining diacritical madks.
5a. Is a repertoire including character names provided?
Yes.
5b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?
Yes.
5c. Are the character shapes attached in a legible form suitable for review?
Yes.
6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson.
6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, Fontographer.
7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
No.
7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
No.
8. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?
No.
9. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.
To be done.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
No.
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
Yes.
2b. If YES, with whom?
David Solnit.
2c. If YES, available relevant documents
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
Kayah Li is used in Myanmar and Thailand.
4a. The context of use for the proposed characters (type of use; common or rare)
Used to write the Kayah Li language.
4b. Reference
5a. Are the proposed characters in current use by the user community?
Yes.
5b. If YES, where?
In Myanmar and Thailand, and in diaspora.
6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?
Yes. Positions AB80-ABAF are proposed.

6b. If YES, is a rationale provided?
Yes.
6c. If YES, reference
Contemporary use and accordance with the Roadmap.
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
Yes.
8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
No.
8b. If YES, is a rationale for its inclusion provided?
8c. If YES, reference
9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
No.
9b. If YES, is a rationale for its inclusion provided?
9c. If YES, reference
10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
No.
10b. If YES, is a rationale for its inclusion provided?
10c. If YES, reference
11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?
Yes.
11b. If YES, is a rationale for such use provided?
No.
11c. If YES, reference
12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
No.
12b. If YES, reference
13a. Does the proposal contain characters with any special properties such as control function or similar semantics?
No.
13b. If YES, describe in detail (include attachment if necessary)
14a. Does the proposal contain any Ideographic compatibility character(s)?
No.
14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

TABLE XXX - Row A9: KAYAH LI

	A90	A91	A92
0	⃞	⃞	⃞
1	⃞	⃞	⃞
2	V	⃞	⃞
3	⃞	⃞	⃞
4	⃞	⃞	⃞
5	⌐	⃞	⃞
6	⃞	⃞	⃞
7	⃞	⃞	⃞
8	⌐	⃞	⃞
9	⃞	⃞	⃞
A	⃞	⃞	⃞
B	⃞	⃞	⃞
C	⃞	⃞	⃞
D	⃞	⃞	⃞
E	⃞	⃞	
F	⃞	⃞	

G = 00
P = 00

TABLE XX - Row A9: KAYAH LI

hex	Name	hex	Name
00	KAYAH LI DIGIT ZERO		
01	KAYAH LI DIGIT ONE		
02	KAYAH LI DIGIT TWO		
03	KAYAH LI DIGIT THREE		
04	KAYAH LI DIGIT FOUR		
05	KAYAH LI DIGIT FIVE		
06	KAYAH LI DIGIT SIX		
07	KAYAH LI DIGIT SEVEN		
08	KAYAH LI DIGIT EIGHT		
09	KAYAH LI DIGIT NINE		
0A	KAYAH LI LETTER KA		
0B	KAYAH LI LETTER KHA		
0C	KAYAH LI LETTER GA		
0D	KAYAH LI LETTER NGA		
0E	KAYAH LI LETTER SA		
0F	KAYAH LI LETTER SHA		
10	KAYAH LI LETTER ZHA		
11	KAYAH LI LETTER NYA		
12	KAYAH LI LETTER TA		
13	KAYAH LI LETTER THA		
14	KAYAH LI LETTER NA		
15	KAYAH LI LETTER PA		
16	KAYAH LI LETTER PHA		
17	KAYAH LI LETTER MA		
18	KAYAH LI LETTER DA		
19	KAYAH LI LETTER BA		
1A	KAYAH LI LETTER RA		
1B	KAYAH LI LETTER YA		
1C	KAYAH LI LETTER LA		
1D	KAYAH LI LETTER WA		
1E	KAYAH LI LETTER THHA		
1F	KAYAH LI LETTER HA		
20	KAYAH LI LETTER VA		
21	KAYAH LI LETTER CA		
22	KAYAH LI LETTER A		
23	KAYAH LI LETTER AE		
24	KAYAH LI LETTER I		
25	KAYAH LI LETTER OO		
26	KAYAH LI VOWEL SIGN Y		
27	KAYAH LI VOWEL SIGN E		
28	KAYAH LI VOWEL SIGN U		
29	KAYAH LI VOWEL SIGN EE		
2A	KAYAH LI VOWEL SIGN O		
2B	KAYAH LI HIGH TONE		
2C	KAYAH LI LOW TONE		
2D	KAYAH LI MID TONE		
2E	(This position shall not be used)		
2F	(This position shall not be used)		