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 Rejang 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 Rejang script in the BMP of the UCS.

Rejang is spoken by a million people living in Indonesia on the island of Sumatra in the southwest highlands, north Bengkulu Province, around Argamakmur, Muaraaman, Curuo, and Kapahiang. Most of its users live in fairly remote rural areas, of whom slightly less than half of them are literate. The traditional Rejang corpus consists chiefly of ritual texts, medical incantations, and poetry.

Structure
Vowel signs are used in a manner similar to that employed by other Brahmi-derived scripts. Consonants have an inherent /a/ vowel sound.

Ordering
An ordering based on the traditional order of the Brahmic script, though with some differences. The Javanese *ha na ca ra ka* ordering is not used.

A. Administrative
1. Title
Proposal for encoding the Rejang 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
Rejang.
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

37

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 2

4b. Is a rationale provided for the choice?

Yes.

4c. If YES, reference

Rejang requires Level 2 implementation as other Brahmic scripts do.

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. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database http://www.unicode.org/Public/UNIDATA/ UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

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.)?

No.

2b. If YES, with whom?

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?

Rejang is used on the island of Bali in Indonesia.

4a. The context of use for the proposed characters (type of use; common or rare)

Used to write the Rejang language.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

In Bali.

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 1B00-1B7F 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?

Yes. The "nuktated" extensions for Arabic can all be composed of base characters and the proposed BALINESE SIGN NUKTA.

9b. If YES, is a rationale for its inclusion provided?

No. This is a preliminary proposal. One question which would need to be dealt with is the interaction of the nukta with the other vowel matras. Another is the repertoire: Rejang seems to have six nuktated consonants for Arabic, but Javanese appears to have a repertoire of eight.

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?

Yes.

11c. If YES, reference

Brahmic vowels.

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 XX - Row A9: REJANG

	A03	A94	A95
0	𑜈	𑜈	○
1	𑜈	𑜈	○
2	𑜈	𑜈	○
3	𑜈	𑜈	○
4	𑜈	𑜈	𑜈
5	𑜈	𑜈	
6	𑜈	𑜈	
7	/	○	
8	𑜈	○	
9	𑜈	○	
A	𑜈	○	
B	𑜈	○	
C	𑜈	○	
D	𑜈	○	
E	𑜈	○	
F	𑜈	○	

G = 00
P = 00

4

TABLE XX - Row A9: REJANG

hex	Name	hex	Name
30	REJANG LETTER KA		
31	REJANG LETTER GA		
32	REJANG LETTER NGA		
33	REJANG LETTER TA		
34	REJANG LETTER DA		
35	REJANG LETTER NA		
36	REJANG LETTER PA		
37	REJANG LETTER BA		
38	REJANG LETTER MA		
39	REJANG LETTER CA		
3A	REJANG LETTER JA		
3B	REJANG LETTER NYA		
3C	REJANG LETTER SA		
3D	REJANG LETTER RA		
3E	REJANG LETTER LA		
3F	REJANG LETTER YA		
40	REJANG LETTER WA		
41	REJANG LETTER HA		
42	REJANG LETTER MBA		
43	REJANG LETTER NGGA		
44	REJANG LETTER NDA		
45	REJANG LETTER NYJA		
46	REJANG LETTER A		
47	REJANG KATULANG (ng)		
48	REJANG DUO DEATAS (n)		
49	REJANG KAJUNJUNG (r)		
4A	REJANG VISARGA		
4B	REJANG VOWEL SIGN KALUAN (i)		
4C	REJANG VOWEL SIGN KAMITAN (u)		
4D	REJANG VOWEL SIGN KAMICA (e)		
4E	REJANG VOWEL SIGN KATILING (ai)		
4F	REJANG VOWEL SIGN O		
50	REJANG VOWEL SIGN KATULUNG (au)		
51	REJANG VOWEL SIGN EU		
52	REJANG VOWEL SIGN KAJINA (ea)		
53	REJANG VIRAMA		
54	REJANG SECTION MARK		
55	(This position shall not be used)		
56	(This position shall not be used)		
57	(This position shall not be used)		
58	(This position shall not be used)		
59	(This position shall not be used)		
5A	(This position shall not be used)		
5B	(This position shall not be used)		
5C	(This position shall not be used)		
5D	(This position shall not be used)		
5E	(This position shall not be used)		
5F	(This position shall not be used)		