DATE: 2002-02-13

DOC TYPE:	Expert contribution
TITLE:	**Proposal to deprecate Khmer characters**
SOURCE:	CHEA Sok Huor, LAO Kim Leang, HARADA Shiro, Norbert KLEIN
PROJECT:	
STATUS:	Proposal
ACTION ID:	
DUE DATE:	--
DISTRIBUTION:	Worldwide
MEDIUM:	Paper
NO. OF PAGES:	

A. Administrative

1. Title	Proposal to deprecate Khmer characters
2. Requester's name	CHEA Sok Huor, LAO Kim Leang, HARADA Shiro, Norbert KLEIN
3. Requester type	Expert request.
4. Submission date	2002-02-13
5. Requester's reference	ISO/IETC JTC 1/SC 2/WG 2 N2380R
6a. Completion	This is a complete proposal.
6b. More information to be provided?	Only as required for clarification.

B. Technical – General

1a. New script? Name?	No.
1b. Addition of characters to existing block? Name?	No
2. Number of characters	6
3. Proposed category	
4. Proposed level of implementation and rationale	
5a. Character names included in proposal?	N/A
5b. Character names in accordance with guidelines?	N/A

5c. Character shapes reviewable?	
6a. Who will provide computerized font?	
6b. Font currently available?	
6c. Font format?	TrueType.
7a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?	
7b. Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?	
8. Does the proposal address other aspects of character data processing?	Yes. This proposal suggests the deprication of characters defined in Unicode 3.1

C. Technical – Justification

1. Contact with the user community?	Yes. National Body Contribution
2. Information on the user community?	Native.
3a. The context of use for the proposed characters?	
3b. Reference	
4a. Proposed characters in current use?	No.
4b. Where?	
5a. Characters should be encoded entirely in BMP?	
5b. Rationale	The characters to be deprecated are not valid for use with Cambodian
6. Should characters be kept in a continuous range?	
7a. Can the characters be considered a presentation form of an existing character or character sequence?	No
7b. Where?	
7c. Reference	
8a. Can any of the characters be considered to be similar (in appearance or function) to an existing character?	

8b. Where?	
8c. Reference	
9a. Combining characters or use of composite sequences included?	
9b. List of composite sequences and their corresponding glyph images provided?	
10. Characters with any special properties such as control function, etc. included?	

D. SC2/WG2 Administrative

To be completed by SC2/WG2

1. Relevant SC 2/WG 2 document numbers:	
2. Status (list of meeting number and corresponding action or disposition)	
3. Additional contact to user communities, liaison organizations etc.	
4. Assigned category and assigned priority/time frame	
Other Comments	

This proposal suggests that the following characters be deprecated from use with Unicode 3.2.

17A3 – This character does not actually exist in Khmer. According to the Unicode Standard, this is used for transliteration of Pali/Sanskrit words. However, it is not enough to include this because it can be represented by the consonant 17A2.

17A4 – This character does not actually exist in Khmer. According to the Unicode Standard, this is used for transliteration of Pali/Sanskrit words. This character is actually a ligature that can be represented by 17A2 + 17B6.

17B4 and 17B5 – These are two inherent vowels that have never been used in Khmer and do not actually exist.

17D3 – This is presumably included to represent the first August of leap year in the lunar calendar. However, we cannot find any code point assigned for the second August. We have submitted a separate proposal to handle Khmer Lunar Dates.

17D8 – This is an abbreviation for the Khmer word meaning "et cetera" expressed like "etc." There is no need for a special code point to represent this. Furthermore, there are other ways of abbreviating this word and it would be inconsistent to include only one of them in the character code table.

In addition to deprecating the above characters, the Cambodian body would like to have annotations added for the following two characters.

17A8 – This character is a ligature of 17A7 + 1780. The recommended usage is to type the two characters instead of the ligature form. The weight for sorting should be the same for these two representations.

17B2 – This character is an alternate form of 17B1. The recommended usage is to use the 17B1. The weight for sorting should be the same for these two representations.