

**ISO/IEC JTC1/SC2/WG2 N1642****DATE:** 1997-09-18

DOC TYPE: Expert contribution
TITLE: **Proposal to encode Cirth in Plane 1 of ISO/IEC 10646-2**
SOURCE: Michael Everson, EGT (IE)
PROJECT: JTC1.02.18.02
STATUS: Proposal.
ACTION ID: FYI
DUE DATE: --
DISTRIBUTION: Worldwide
MEDIUM: Paper and web
NO. OF PAGES: 5

A. Administrative

1. Title	Proposal to encode Cirth in Plane 1 of ISO/IEC 10646-2
2. Requester's name	Michael Everson
3. Requester type	Expert request
4. Submission date	1997-09-18
5. Requester's reference	
6a. Completion	This is a complete proposal.
6b. More information to be provided?	No

B. Technical -- General

1a. New script? Name?	Yes. Cirth
1b. Addition of characters to existing block? Name?	No.
2. Number of characters	103
3. Proposed category	Category B.1
4. Proposed level of implementation and rationale	Level 3
5a. Character names included in proposal?	Yes
5b. Character names in accordance with guidelines?	Yes
5c. Character shapes reviewable?	Yes
6a. Who will provide computerized font?	Michael Everson, Everson Gunn Teoranta
6b. Font currently available?	Michael Everson, Everson Gunn Teoranta
6c. Font format?	TrueType
7a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?	Yes.
7b. Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?	No
8. Does the proposal address other aspects of character data processing?	No

C. Technical -- Justification	
1. Contact with the user community?	Yes. There are several Internet discussion lists and web sites.
2. Information on the user community?	The Cirth enjoy both scholarly and popular use.
3a. The context of use for the proposed characters?	Used to write Khuzdul, Sindarin, English, and other languages.
3b. Reference	
4a. Proposed characters in current use?	Yes
4b. Where?	By scholars and enthusiasts.
5a. Characters should be encoded entirely in BMP?	No. Positions U+0001 3380 - U+0001 337F are proposed for the encoding.
5b. Rationale	Accordance with the Roadmap.
6. Should characters be kept in a continuous range?	Yes
7a. Can the characters be considered a presentation form of an existing character or character sequence?	Some of the Cirth look like Germanic runes but they are considerably more numerous, and the phonetic systems used with them are entirely different.
7b. Where?	
7c. Reference	
8a. Can any of the characters be considered to be similar (in appearance or function) to an existing character?	One could say that the punctuation characters look like other characters, but it is advisable to encode them because they are used in a relatively systematic way in the writing system and it keeps things tidier.
8b. Where?	
8c. Reference	
9a. Combining characters or use of composite sequences included?	Yes, the last three characters are combining characters.
9b. List of composite sequences and their corresponding glyph images provided?	No
10. Characters with any special properties such as control function, etc. included?	No
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	

The Cirth script was invented by the philologist and author J. R. R. Tolkien as part of the mythological world he created and was widely popularized through his work, *The Lord of the Rings*, *The Silmarillion*, etc. Along with a family of artificial languages and a large corpus of etymological data describing their relationships, the Cirth script has attracted the attention of a large community of linguists and other enthusiasts interested in this expression of Tolkien's expertise in historical and comparative linguistics. It can be categorized as a Category D (Attested Extinct) alphabet: there is a relatively limited corpus, and a relatively small (but existent) scholarly body studying it. In order to set a standard Cirth character coding for such scholars and enthusiasts, it has been suggested that this character set be included into the Unicode standard and ISO 10646.

8 columns are reserved to encode the Cirth. The last column is currently unused, and is reserved for future discoveries in the Tolkien manuscripts. The Cirth was and is used to write the languages Quenya, Sindarin, and Khuzdul. It has also been used to write English, as on the title page of *The Lord Of The Rings*.

General Principles of the Cirth script

The Cirth are a Runic-type alphabet, although they are not connected with Nordic runes except due to a general resemblance resulting from the constraints of letterforms carved in wood or stone. Some of the Cirth had two different forms, which seem to represent glyphic variants. The Cirth were written from left to right. No positional variants or non-spacing marks exist.

Ordering follows the presentation of the Eregion and Moria Cirth and the earliest Beleriand runes (from *The Return of the King*, Appendix E, and *The Treason of Isengard*, Appendix on Runes). Additional Cirth from Doriath and Noldor have been inserted into this order, as have other Cirth used for English, etc. Where duplication in letter names occurs, a modifier has been added to the name to differentiate it from the primary form. Pronounceable or meaningful names are not known for the Cirth, so their phonetic values are given in the names. Long vowels are written doubled.

Punctuation

Little is known about punctuation marks, though four have been identified: a single dot serves sometimes to separate letters or words; two vertical dots is used to break up groups longer than a word; three or four vertical dots are used at the beginning and ending of texts. Only three Cirth digits are extant; each is formed by placing a dot beneath an existing Certh, so that non-spacing dot has been encoded here.

Sometimes word space is not used; word separation may be achieved in that case with U+200B, ZERO WIDTH SPACE. Hyphenation is not used; words may be broken after any LETTER.

U+xx80	CIRTH LETTER P
U+xx81	CIRTH LETTER B
U+xx82	CIRTH LETTER F
U+xx83	CIRTH LETTER V
U+xx84	CIRTH LETTER HW
U+xx85	CIRTH LETTER M
U+xx86	CIRTH LETTER MB
U+xx87	CIRTH LETTER SP
U+xx88	CIRTH LETTER SB
U+xx89	CIRTH LETTER SC
U+xx8A	CIRTH LETTER SG
U+xx8B	CIRTH LETTER T
U+xx8C	CIRTH LETTER D
U+xx8D	CIRTH LETTER TH
U+xx8E	CIRTH LETTER DH
U+xx8F	CIRTH LETTER N
U+xx90	CIRTH LETTER NDZH
U+xx91	CIRTH LETTER DORIAN KW
U+xx92	CIRTH LETTER DORIAN GW
U+xx93	CIRTH LETTER DORIAN KHW
U+xx94	CIRTH LETTER DORIAN GHW
U+xx95	CIRTH LETTER DORIAN L
U+xx96	CIRTH LETTER ENGLISH ND
U+xx97	CIRTH LETTER CH
U+xx98	CIRTH LETTER J
U+xx99	CIRTH LETTER SH
U+xx9A	CIRTH LETTER ZH
U+xx9B	CIRTH LETTER NJ
U+xx9C	CIRTH LETTER K
U+xx9D	CIRTH LETTER G
U+xx9E	CIRTH LETTER KH
U+xx9F	CIRTH LETTER GH
U+xxA0	CIRTH LETTER ENG
U+xxA1	CIRTH LETTER KW
U+xxA2	CIRTH LETTER GW
U+xxA3	CIRTH LETTER KHW
U+xxA4	CIRTH LETTER GHW
U+xxA5	CIRTH LETTER NGW
U+xxA6	CIRTH LETTER NW
U+xxA7	CIRTH LETTER DORIAN Z
U+xxA8	CIRTH LETTER R
U+xxA9	CIRTH LETTER RH
U+xxAA	CIRTH LETTER L
U+xxAB	CIRTH LETTER LH
U+xxAC	CIRTH LETTER NG
U+xxAD	CIRTH LETTER S
U+xxAE	CIRTH LETTER KHUZDUL GLOTTAL STOP
U+xxAF	CIRTH LETTER Z
U+xxB0	CIRTH LETTER KHUZDUL NG
U+xxB1	CIRTH LETTER ND
U+xxB2	CIRTH LETTER EI
U+xxB3	CIRTH LETTER IU

```

U+xxB4 CIRTH LETTER I
U+xxB5 CIRTH LETTER KHUZDUL Y
U+xxB6 CIRTH LETTER KHUZDUL HY
U+xxB7 CIRTH LETTER U
U+xxB8 CIRTH LETTER UU
U+xxB9 CIRTH LETTER W
U+xxBA CIRTH LETTER UE
U+xxBB CIRTH LETTER UI
U+xxBC CIRTH LETTER E
U+xxBD CIRTH LETTER EE
U+xxBE CIRTH LETTER A
U+xxBF CIRTH LETTER AA
U+xxC0 CIRTH LETTER AI
U+xxC1 CIRTH LETTER AU
U+xxC2 CIRTH LETTER AY
U+xxC3 CIRTH LETTER AE
U+xxC4 CIRTH LETTER EA
U+xxC5 CIRTH LETTER EW
U+xxC6 CIRTH LETTER O
U+xxC7 CIRTH LETTER OO
U+xxC8 CIRTH LETTER OE
U+xxC9 CIRTH LETTER NOLDORIAN O
U+xxCA CIRTH LETTER NOLDORIAN OO
U+xxCB CIRTH LETTER IO
U+xxCC CIRTH LETTER EU
U+xxCD CIRTH LETTER OU
U+xxCE CIRTH LETTER NOLDORIAN OE
U+xxCF CIRTH LETTER KHUZDUL N
U+xxD0 CIRTH LETTER H
U+xxD1 CIRTH LETTER KHUZDUL LEFT-POINTING SCHWA
U+xxD2 CIRTH LETTER KHUZDUL RIGHT-POINTING SCHWA
U+xxD3 CIRTH LETTER DORIAN O
U+xxD4 CIRTH LETTER KHUZDUL PS
U+xxD5 CIRTH LETTER KHUZDUL TS
U+xxD6 CIRTH MODIFIER LETTER H
U+xxD7 CIRTH ENGLISH THE
U+xxD8 CIRTH AMPERSAND
U+xxD9 CIRTH NOLDORIAN L
U+xxDA CIRTH ENGLISH OF
U+xxDB CIRTH LETTER Y
U+xxDC CIRTH LETTER VARIANT Y
U+xxDD CIRTH LETTER YY
U+xxDE CIRTH LETTER NOLDORIAN OOE
U+xxDF CIRTH LETTER NOLDORIAN OE
U+xxE0 CIRTH SEPARATOR SINGLE DOT
U+xxE1 CIRTH SEPARATOR DOUBLE DOT
U+xxE2 CIRTH SEPARATOR TRIPLE DOT
U+xxE3 CIRTH START OR END OF TEXT
U+xxE4 CIRTH SEPARATOR DOUBLE PIPE
U+xxE5 CIRTH COMBINING NASAL MARK
U+xxE6 CIRTH COMBINING LENGTH MARK
U+xxE7 CIRTH NUMERIC DOT
U+xxE8 (This position shall not be used)
U+xxE9 (This position shall not be used)
U+xxEA (This position shall not be used)
U+xxEB (This position shall not be used)
U+xxEC (This position shall not be used)
U+xxED (This position shall not be used)
U+xxEE (This position shall not be used)
U+xxEF (This position shall not be used)
U+xxF0 (This position shall not be used)
U+xxF1 (This position shall not be used)

```

U+xxF2 (This position shall not be used)
 U+xxF3 (This position shall not be used)
 U+xxF4 (This position shall not be used)
 U+xxF5 (This position shall not be used)
 U+xxF6 (This position shall not be used)
 U+xxF7 (This position shall not be used)
 U+xxF8 (This position shall not be used)
 U+xxF9 (This position shall not be used)
 U+xxFA (This position shall not be used)
 U+xxFB (This position shall not be used)
 U+xxFC (This position shall not be used)
 U+xxFD (This position shall not be used)
 U+xxFE (This position shall not be used)
 U+xxFF (This position shall not be used)

	E08	E09	E0A	E0B	E0C	E0D	E0E	E0F
0	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ	·	
1	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ	:	
2	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ	:	
3	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ	:	
4	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		
5	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ	⊙	
6	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ	⊙	
7	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ	⊙	
8	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		
9	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		
A	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		
B	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		
C	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		
D	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		
E	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		
F	ƒ	ƒ	ƒ	ƒ	ƒ	ƒ		