



ISO/IEC JTC1/SC2/WG2 N1689

DATE: 1998-01-18

ISO/IEC JTC1/SC2/WG2
 Universal Multiple-Octet Coded Character Set (UCS) - ISO/IEC 10646
 Secretariat: ANSI

DOC TYPE: Expert contribution
TITLE: Proposal to encode South Arabian 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: 3 (printed at 80%)

A. Administrative

1. Title	Proposal to encode South Arabian in Plane 1 of ISO/IEC 10646-2
2. Requester's name	Michael Everson
3. Requester type	Expert request
4. Submission date	1998-01-18
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. South Arabian
1b. Addition of characters to existing block? Name?	No.
2. Number of characters	29
3. Proposed category	Category D
4. Proposed level of implementation and rationale	Level 1
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
6b. Font currently available?	Michael Everson
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?	Yes

C. Technical -- Justification

1. Contact with the user community?	No.
2. Information on the user community?	South Arabian enjoys mostly scholarly use.
3a. The context of use for the proposed characters?	Used to represent texts from the Sabaeen and Minaean kingdoms in Southern Arabia, ca. 500 BCE - 600 CE.
3b. Reference	See below.
4a. Proposed characters in current use?	Yes.
4b. Where?	By scholars.
5a. Characters should be encoded entirely in BMP?	No. Plane 1
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?	No.
7b. Where?	
7c. Reference	
8a. Can any of the characters be considered to be similar (in appearance or function) to an existing character?	No
8b. Where?	
8c. Reference	
9a. Combining characters or use of composite sequences included?	No.
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 script known as South Arabian is related to the Sinaitic and early Semitic alphabets, but its shapes are remarkably unique for such a derivation. It is also an ancestor of the modern Ethiopic script. Inscriptions in this script are found in southern Arabia (the ancient Sabaeen and Minaean kingdoms)

dating from as far back as 500 BCE. The script was apparently used until about 600 CE. The alphabetic order is as given in Healey 1990, based on fragmentary evidence. This proposal is, apart from the names, identical to the proposal of Rick McGowan in UTR #3. The South Arabian default directionality is RTL though some boustrophedon inscriptions are found.

Issues:

Faulmann gives a set of numbers. This needs to be looked into with more modern sources and experts. Is there any punctuation that needs to be coded?

The names given here are versions of their Latin transliterations, using the same conventions as for Ethiopic where possible. Do actual names exist for these characters?

U+0001xx00	SOUTH ARABIAN LETTER H		xx0	xx1
U+0001xx01	SOUTH ARABIAN LETTER L			
U+0001xx02	SOUTH ARABIAN LETTER HH	0	ሀ	ሐ
U+0001xx03	SOUTH ARABIAN LETTER M			
U+0001xx04	SOUTH ARABIAN LETTER Q			
U+0001xx05	SOUTH ARABIAN LETTER W			
U+0001xx06	SOUTH ARABIAN LETTER S	1	ሀ	ዐ
U+0001xx07	SOUTH ARABIAN LETTER R			
U+0001xx08	SOUTH ARABIAN LETTER B			
U+0001xx09	SOUTH ARABIAN LETTER T	2	ሀ	ፀ
U+0001xx0A	SOUTH ARABIAN LETTER SH			
U+0001xx0B	SOUTH ARABIAN LETTER K			
U+0001xx0C	SOUTH ARABIAN LETTER N	3	ሀ	ገ
U+0001xx0D	SOUTH ARABIAN LETTER X			
U+0001xx0E	SOUTH ARABIAN LETTER SZ			
U+0001xx0F	SOUTH ARABIAN LETTER F	4	ሀ	ቀ
U+0001xx10	SOUTH ARABIAN LETTER HAMZA			
U+0001xx11	SOUTH ARABIAN LETTER AYIN	5	ፀ	ፐ
U+0001xx12	SOUTH ARABIAN LETTER TZ			
U+0001xx13	SOUTH ARABIAN LETTER G			
U+0001xx14	SOUTH ARABIAN LETTER D	6	ሀ	፱
U+0001xx15	SOUTH ARABIAN LETTER J			
U+0001xx16	SOUTH ARABIAN LETTER TH			
U+0001xx17	SOUTH ARABIAN LETTER Z	7	ሀ	፳
U+0001xx18	SOUTH ARABIAN LETTER DJ			
U+0001xx19	SOUTH ARABIAN LETTER Y			
U+0001xx1A	SOUTH ARABIAN LETTER TJ	8	ሀ	ዘ
U+0001xx1B	SOUTH ARABIAN LETTER TS			
U+0001xx1C	SOUTH ARABIAN LETTER ZZ			
U+0001xx1D	(This position shall not be used)	9	ሀ	ዓ
U+0001xx1E	(This position shall not be used)			
U+0001xx1F	(This position shall not be used)	A	ሀ	፳
		B	ሀ	፳
		C	ሀ	ዓ
		D	ሀ	
		E	ሀ	
		F	ሀ	

Bibliography

Faulmann, Carl. 1990 (1880). *Das Buch der Schrift*. Frankfurt am Main: Eichborn. ISBN 3-8218-1720-8

Haarmann, Harald. 1990. *Universalgeschichte der Schrift*. Frankfurt/Main; New York: Campus. ISBN 3-593-34346-0

Healey, John F. 1990. *The early alphabet*. (Reading the Past) Trustees of the British Museum: British Museum Press. ISBN 0-7141-8073-4

Unicode Consortium. 1992. *Unicode Technical Report #3: exploratory proposals*

HTML Michael Everson, everson@indigo.ie, <http://www.indigo.ie/egt>, Dublin, 1998-01-18