

Universal Multiple-Octet Coded Character Set  
International Organization for Standardization  
Международная организация по стандартизации

**Doc Type: Working Group Document****Title: On encoding the Old Hungarian rovásírás in the UCS****Source: Michael Everson****Status: Expert Contribution****Action: For consideration by JTC1/SC2/WG2****Date: 1998-05-02**

Since the circulation of the first proposal to include Old Hungarian in the Universal Character Set, feedback has been given from two major, slightly overlapping, sources. The first was the ROVAS@NIMRUD.EET.BME.HU e-mail discussion list, which exists for the discussion of the *rovásírás* writing system and other matters of Hungarian linguistics. (Discussion on that list is chiefly in Hungarian.) The other source for feedback is the Old Hungarian Culture Association, which has evaluated a set of existing fonts and provided a definitive Old Hungarian font containing the basic alphabet and a number of “standard ligatures”.

In my original proposal, I pointed out that there were two main issues regarding the encoding of Old Hungarian:

- 1) The treatment of the ligatures in the UCS
- 2) The default directionality in the UCS

The original source font I received from the Rovás Honlap web site contained 64 ligatures, shown in positions 30-6F of N1686. In other sources listed in N1686 I found 71 additional ligatures shown in positions 70-B7 of N1686. After reviewing these, the Old Hungarian Culture Association has settled upon a standard number of 88 ligatures (“the others are not safe”). The remaining “safe” ligatures, and a monogram I have seen on the internet, however, underscore the fact that, while there are “standard” ligatures, ligature formation in Old Hungarian is *productive* – and, again, as can be seen from an abundance of examples on the internet, ligature formation is *optional*. I therefore propose that in the UCS, *only* the basic set of letters be encoded and that ZERO WIDTH JOINER be used in any instance where a ligature is required, whether the ligature be a standard ligature like Ꞓ GA (Ꞓ + ꞑ) or Ꞓ CSIN (Ꞓ + ꞑ + Ꞓ), or a nonce ligature like Ꞓ OSKÁNT (Ꞓ + Ꞓ + ꞑ + Ꞓ (= Ꞓ + ꞑ + Ꞓ)). One advantage of this is that the orthography of words written in Old Hungarian would be consistent, facilitating lexical searching of Old Hungarian texts.

The default directionality for Old Hungarian is less simple to decide, though for my part I have come down firmly on the side of supporting its original RTL directionality as the default. From the time of its original use to the present day, Old Hungarian can be written either RTL or LTR; in earlier times RTL predominated. The Old Hungarian Culture Association strongly favours a default of LTR, because it is practical for implementation on 8-bit PCs, and they are used to that. A strong opposition, however, was evident in the ROVAS e-mail list, from other ordinary users of the script.

Use of Old Hungarian is not confined to computers. Apparently, though admittedly not a subject universally enjoyed, Old Hungarian is known and used by ordinary Hungarian schoolchildren, and has been for a long time – certainly since before the use of the script on computers. A number of people on the ROVAS list admitted to having learned Old Hungarian script in school – for fun, for secrecy, and because it was Hungarian – and when questioned about directionality, all stated that the RTL directionality was an important feature of the script to them. Such statements from ordinary users – as well as the fairly large number of RTL gifs available on numerous websites, suggests to me that RTL directionality should be the default for Old Hungarian, and that directional overrides be used to effect LTR presentation.

I am a little uncomfortable coming down on the side of the non-specialist, because I’m a fairly academic type myself. I understand what Hosszú Gábor has said: “It is a pity that the RTL Party is louder, because the LTR party is more productive...” In this instance I am guided by our experience with Etruscan, which also had RTL and LTR directionality during its heyday. Today, however, Etruscan enjoys almost *no* popular use, excepting scholarly use, and scholars are unanimous in preferring LTR.

For Old Hungarian, however, the impression I get is that the scholars are used to LTR because it has been practical to implement on computers and is less troublesome to read for people used to reading the Latin script – but it seems that those particular needs should be met with the directional overrides, and that the native directionality, supported very strongly by ordinary users for cultural reasons, should be respected in this case. Ease of implementation on 8-bit PCs should not determine the default for this script.

With regard to the basic repertoire, the Old Hungarian Culture Association has added one character to the proposal in N1686, Ǻ CLOSED E. This character is attested in the same sources which provided the 71 additional ligatures. A separator character, which can have a number of glyph representations (one dot or two dots (or more?)), has also been added.

Since the publication of N1686 the ad-hoc Roadmap Committee observed that there was room in the notional RTL block of the BMP for Old Hungarian (3 columns), and that, since the script seems to enjoy a fair bit of contemporary computer use, it should be allocated to the BMP and not to Plane 1.

Spelling of the character names remains to be determined. Two options are given below.

Old Hungarian web sites of interest: <http://nimrud.eet.bme.hu/rovas/>; <http://fang.fa.gau.hu/~heves/>

TABLE XXX - Row XX: OLD HUNGARIAN

	xx0	xx1	xx2
0	𐌵	𐌶	𐌷
1	𐌸	𐌹	𐌺
2	𐌻	𐌼	𐌽
3	𐌾	𐌿	𐍀
4	𐍁	𐍂	𐍃
5	𐍄	𐍅	𐍆
6	𐍇	𐍈	𐍉
7	𐍊	𐍋	𐍌
8	𐍍	𐍎	𐍏
9	𐍐	𐍑	𐍒
A	𐍓	𐍔	𐍕
B	𐍖	𐍗	𐍘
C	𐍙	𐍚	𐍛
D	𐍜	𐍝	𐍞
E	𐍟	𐍠	·
F	𐍡	𐍢	

G = 00  
P = 00

**TABLE XXX - Row XX: OLD HUNGARIAN**

dec	hex	Name
000	00	OLD HUNGARIAN LETTER A
001	01	OLD HUNGARIAN LETTER AA
002	02	OLD HUNGARIAN LETTER EB
003	03	OLD HUNGARIAN LETTER EC
004	04	OLD HUNGARIAN LETTER ECS
005	05	OLD HUNGARIAN LETTER ED
006	06	OLD HUNGARIAN LETTER E
007	07	OLD HUNGARIAN LETTER EE
008	08	OLD HUNGARIAN LETTER CLOSE E
009	09	OLD HUNGARIAN LETTER EF
010	0A	OLD HUNGARIAN LETTER EG
011	0B	OLD HUNGARIAN LETTER EGY
012	0C	OLD HUNGARIAN LETTER EH
013	0D	OLD HUNGARIAN LETTER I
014	0E	OLD HUNGARIAN LETTER II
015	0F	OLD HUNGARIAN LETTER EJ
016	10	OLD HUNGARIAN LETTER EK
017	11	OLD HUNGARIAN LETTER AK
018	12	OLD HUNGARIAN LETTER EL
019	13	OLD HUNGARIAN LETTER ELY
020	14	OLD HUNGARIAN LETTER EM
021	15	OLD HUNGARIAN LETTER EN
022	16	OLD HUNGARIAN LETTER ENY
023	17	OLD HUNGARIAN LETTER O
024	18	OLD HUNGARIAN LETTER OO
025	19	OLD HUNGARIAN LETTER OE
026	1A	OLD HUNGARIAN LETTER OEE
027	1B	OLD HUNGARIAN LETTER EP
028	1C	OLD HUNGARIAN LETTER ER
029	1D	OLD HUNGARIAN LETTER ES
030	1E	OLD HUNGARIAN LETTER ESZ
031	1F	OLD HUNGARIAN LETTER ET
032	20	OLD HUNGARIAN LETTER ETY
033	21	OLD HUNGARIAN LETTER U
034	22	OLD HUNGARIAN LETTER UU
035	23	OLD HUNGARIAN LETTER UE
036	24	OLD HUNGARIAN LETTER UEE
037	25	OLD HUNGARIAN LETTER EV
038	26	OLD HUNGARIAN LETTER EZ
039	27	OLD HUNGARIAN LETTER EZS
040	28	OLD HUNGARIAN NUMBER ONE
041	29	OLD HUNGARIAN NUMBER FIVE
042	2A	OLD HUNGARIAN NUMBER TEN
043	2B	OLD HUNGARIAN NUMBER FIFTY
044	2C	OLD HUNGARIAN NUMBER ONE HUNDRED
045	2D	OLD HUNGARIAN NUMBER ONE THOUSAND
046	2E	OLD HUNGARIAN SEPARATOR
047	2F	(This position shall not be used)

Group 00

dec	hex	Name
		OLD HUNGARIAN LETTER A
		OLD HUNGARIAN LETTER AA
		OLD HUNGARIAN LETTER BE
		OLD HUNGARIAN LETTER CE
		OLD HUNGARIAN LETTER CSE
		OLD HUNGARIAN LETTER DE
		OLD HUNGARIAN LETTER E
		OLD HUNGARIAN LETTER EE
		OLD HUNGARIAN LETTER CLOSE E
		OLD HUNGARIAN LETTER EF
		OLD HUNGARIAN LETTER GE
		OLD HUNGARIAN LETTER GYE
		OLD HUNGARIAN LETTER HA
		OLD HUNGARIAN LETTER I
		OLD HUNGARIAN LETTER II
		OLD HUNGARIAN LETTER JE
		OLD HUNGARIAN LETTER KA
		OLD HUNGARIAN LETTER AK ???
		OLD HUNGARIAN LETTER EL
		OLD HUNGARIAN LETTER ELY
		OLD HUNGARIAN LETTER EM
		OLD HUNGARIAN LETTER EN
		OLD HUNGARIAN LETTER ENY
		OLD HUNGARIAN LETTER O
		OLD HUNGARIAN LETTER OO
		OLD HUNGARIAN LETTER OE
		OLD HUNGARIAN LETTER OEE
		OLD HUNGARIAN LETTER PE
		OLD HUNGARIAN LETTER ER
		OLD HUNGARIAN LETTER ES
		OLD HUNGARIAN LETTER ESZ
		OLD HUNGARIAN LETTER TE
		OLD HUNGARIAN LETTER TYE
		OLD HUNGARIAN LETTER U
		OLD HUNGARIAN LETTER UU
		OLD HUNGARIAN LETTER UE
		OLD HUNGARIAN LETTER UEE
		OLD HUNGARIAN LETTER VE
		OLD HUNGARIAN LETTER ZE
		OLD HUNGARIAN LETTER ZSE
		OLD HUNGARIAN NUMBER ONE
		OLD HUNGARIAN NUMBER FIVE
		OLD HUNGARIAN NUMBER TEN
		OLD HUNGARIAN NUMBER FIFTY
		OLD HUNGARIAN NUMBER ONE HUNDRED
		OLD HUNGARIAN NUMBER ONE THOUSAND
		OLD HUNGARIAN SEPARATOR
		(This position shall not be used)

Plane 00

Row XX