L2/23-060

Proposal to Encode a Hungarian Forint Symbol in the Unicode Standard

Nules, Vacek
vnules@gmail.com

29 January 2023
Budapest, Hungary

Revision 2

Preamble

This proposal aims to add a „Ft" character as a currency symbol for the Hungarian forint to the Unicode Standard.

The forint is the currency of Hungary since 1946 and is usually abbreviated „Ft" (U+0046 U+0074), but there are many examples of the „Ft" symbol being used as a single character since at least 1970 (see exhibit 5.1) and in current use on computers (see exhibit 1).

Rationale

Contemporary usage of the „Ft" character is mostly limited to legacy devices, largely because no contemporary OS, apart from FreeDOS, seems to support it, nor does the Unicode Standard. The usual convention on Windows and the World Wide Web to represent the forint is U+0046 U+0074. However, given Hungary's weak economy and low technical development, tens of thousands of legacy devices are seeing day-by-day use and the need for the „Ft" character arises to allow documents produced with these devices to be digitized with an 1:1 mapping. The rationale is similar to the peseta sign ₧, which was also included from legacy codepages and represents a currency which has been out of circulation for over 20 years, unlike the forint, which is the circulating currency of Hungary and owing to Hungary's current political and economical situation is expected to still be for at least another 10 years.

Character to be added

U+20C1 ₁ FORINT SIGN

Unicode Data

```
20C1;FORINT SIGN;Sc;0;ET;;;;;N;;;;;
```

Exhibits of Established Uses of a „Ft" Symbol as a Single Character to Represent the Hungarian Forint

Part One: Examples of Use on Computers

This part attempts to gather evidence on the Ft sign being used on actual computers which are likely to produce files to be interchanged with other computers.

Exhibit 1: FreeDOS

FreeDOS is an open source DOS-compatible operating system[1] which is routinely installed on new PCs sold as of the writing of this document to avoid the cost of a pre-installed commercial operating system.

FreeDOS contains a number of codepages in .CPX files, one of which, named EGA10.CPX, contains a codepage identified as "*57781 – Hungarian*"[2], which is almost identical to what Wikipedia lists as "CWI-2"[3], except for codepoint 159 being redefined from "*ƒ*" to "Ft".

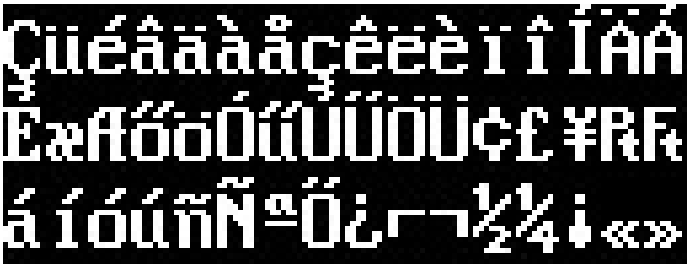

Image 1.1: Partial binary dump of the 8×16 pixel sized font for code page 57781 of the FreeDOS code page file EGA10.CPI which I obtained by extracting EGA10.CPX using an official tool called CPX2CPI.

Image 1.2: Partial binary dump of the 8×14 pixel sized font from the same source

Image 1.3: Partial binary dump of the 8×8 pixel sized font from the same source

Exhibit 2: Primo

PRIMO is a home computer which was developed in the 80's by a [*H*]ungarian computer developer called SZTAKI[4, 5]. PRIMO computers were manufactured from approximately 1984 to 1986, and used a custom codepage which features (a subset of) Hungarian accents and a Ft sign on codepoint 137.

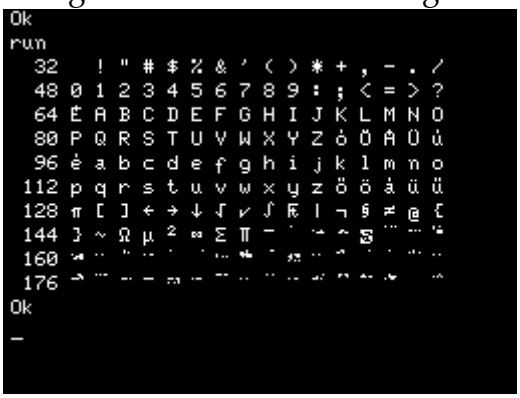

Image 2.1: Character set output in the PC-based Primo emulator *Ultimo*[6], with the Ft sign of the shared-horizontal-bar design featured at codepoint 137, between ∫ and |.

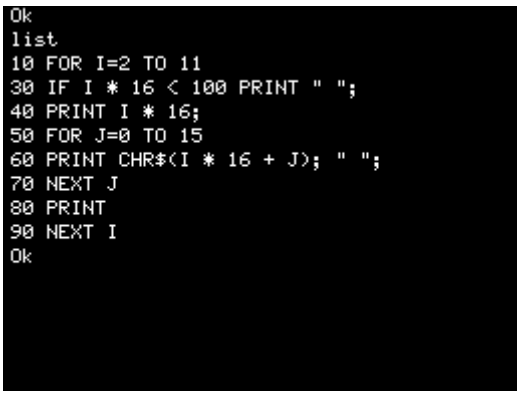

Image 2.2: BASIC source code used to generate Image 2.1, also grabbed from *Ultimo*.

CHR$ kód	karakter	CHR$ kód	karakter	CHR$ kód	karakter
108	l	123	ő	138	|
109	m	124	ö	139	⌐
110	n	125	á	140	§
111	o	126	ü	141	≠
112	p	127	ű	142	@
113	q	128	π	143	⌠
114	r	129	[	144	⌡
115	s	130	]	145	~
116	t	131	←	146	Ω
117	u	132	→	147	µ
118	v	133	↓	148	²
119	w	134	√	149	∞
120	x	135	⌐	150	Σ
121	y	136	∫	151	Π
122	z	137	Ft.		

Image 2.3: Excerpt from the official Primo user's manual[7], showing the character set containing the Ft sign (here including a superfluous period which is not shown in the ROM font)

Exhibit 3: "Homebrew" DOS applications

Since support for the Hungarian accented characters was developed only late in MS-DOS's history and code page 852 contained a lot of superfluous Slavic characters which displaced many of the graphic characters, many software developers in Hungary resorted to drawing their own custom fonts, usually in the form of memory-resident .COM programs and mostly supporting the CWI-2 encoding, of which two will be showcased here.

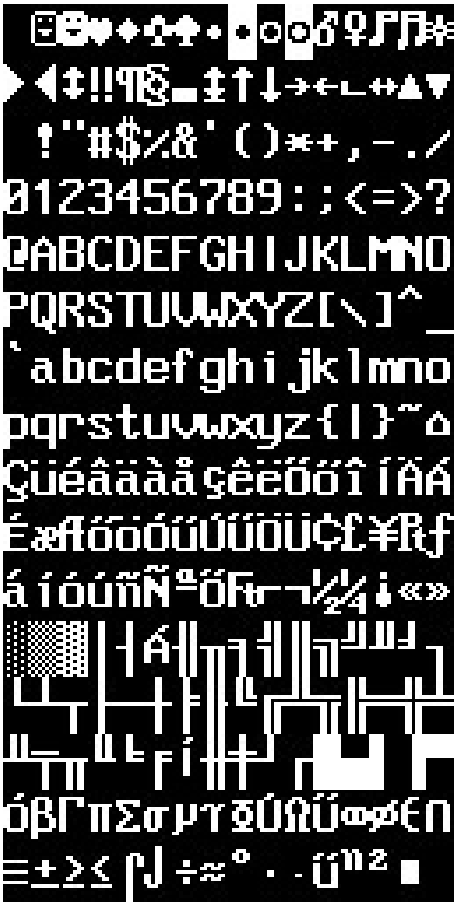

Image 3.1: Complete binary dump of "*gaborfnt.com*", dated April 10, 1995 and found on a salvaged 3½" diskette, a completely documentation-less memory resident font-changing utility (presumably made by someone named *Gábor*) and having correct accents at both the IBM 852 and the CWI-2 codepoints, featuring a Ft sign at codepoint 168.

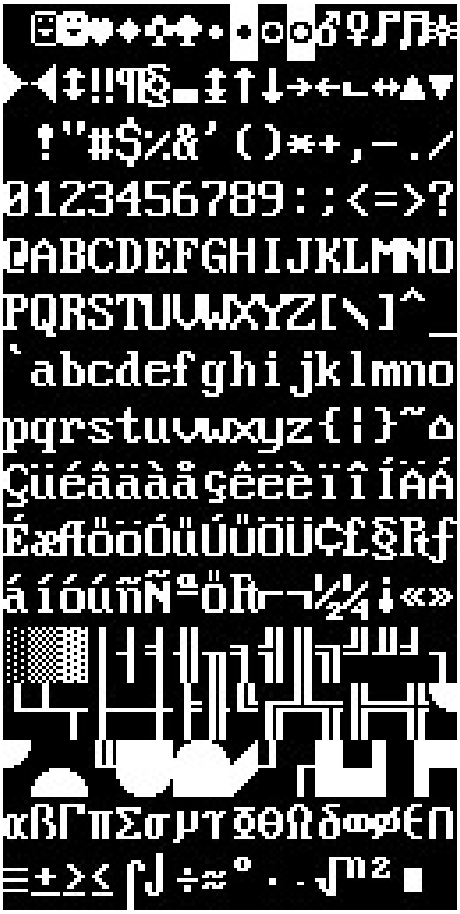

Image 3.2: Complete binary dump of "*font.com*", another memory resident font-changing utility packed in with the proprietary POS software "E-Codirius II". Notice that code points 207–210 and 212–216 contain custom curved and sloped block elements for displaying large digits, which are outside the scope of this proposal. The Ft sign has a shared-horizontal-bar design.

Part Two: Examples of Use in Documents Derived from Computers

This part attempts to gather evidence on the Ft sign being used in printed documents which were made with a computer but the identity of the machine itself could not be deduced yet.

Exhibit 4: TV BASIC book

The book "TV BASIC" was written in 1984 by András Kocsis[8] to accompany a televised, 20-episode educational series on the BASIC programming language[9] broadcasted by the Hungarian State Television. This book contains numerous BASIC listings on pages 317–376, some of which (pictured) contain Ft signs where usual BASIC syntax would require a dollar ($) sign.

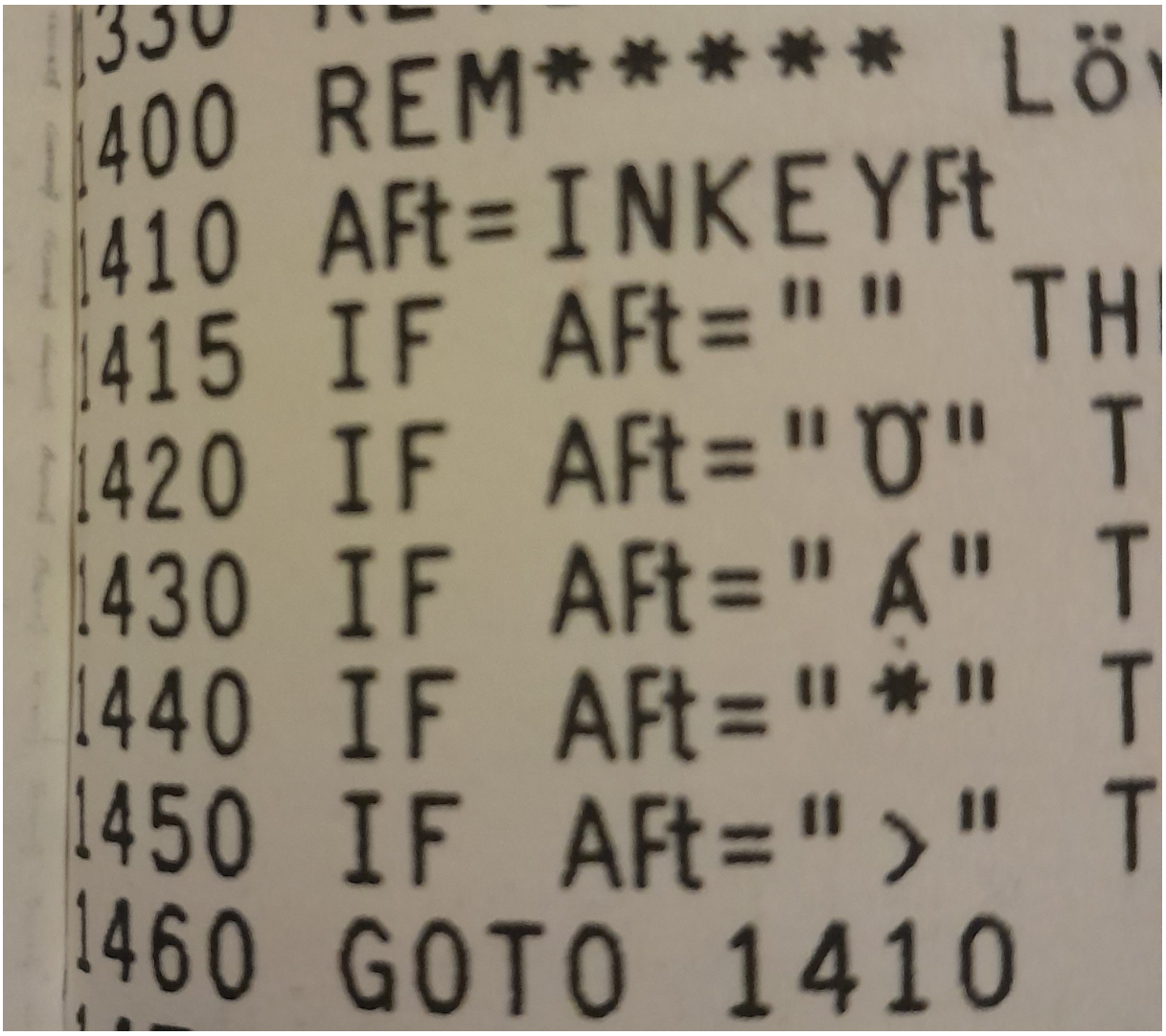

Image 4.1: Photograph of part of page 363 of said book, which features Ft signs on BASIC lines 1410–1450, replacing dollar signs in statements like the variable name A$

and the built-in function INKEY$. Note that some lines contain Hungarian accented letters.

Part Three: Examples of Use in Documents Not Derived from Computers

This part attempts to gather evidence on the Ft sign being used in printed documents which were *not* made with a computer, but whose digitalization at some time may be likely to occur and the need for the Ft sign to be accurately encoded may arise.

Exhibit 5: Typewriters

Image 5.1: Part of the keyboard of a *Robotron 20* typewriter (1970) with a dedicated Ft key.[10]

Image 5.2: Part of the keyboard of an *Optima 26* typewriter (c. 1987) with a dedicated Ft character as a third-level ("alternate") character inputtable by pressing *CODE* and the 4 key.[11]

Part Four: Examples of Use in Areas Not Designed to Interact with Computers

This part attempts to gather evidence on the Ft sign being used on devices which are not expected to interchange character data with computers.

Exhibit 6: Public payphone

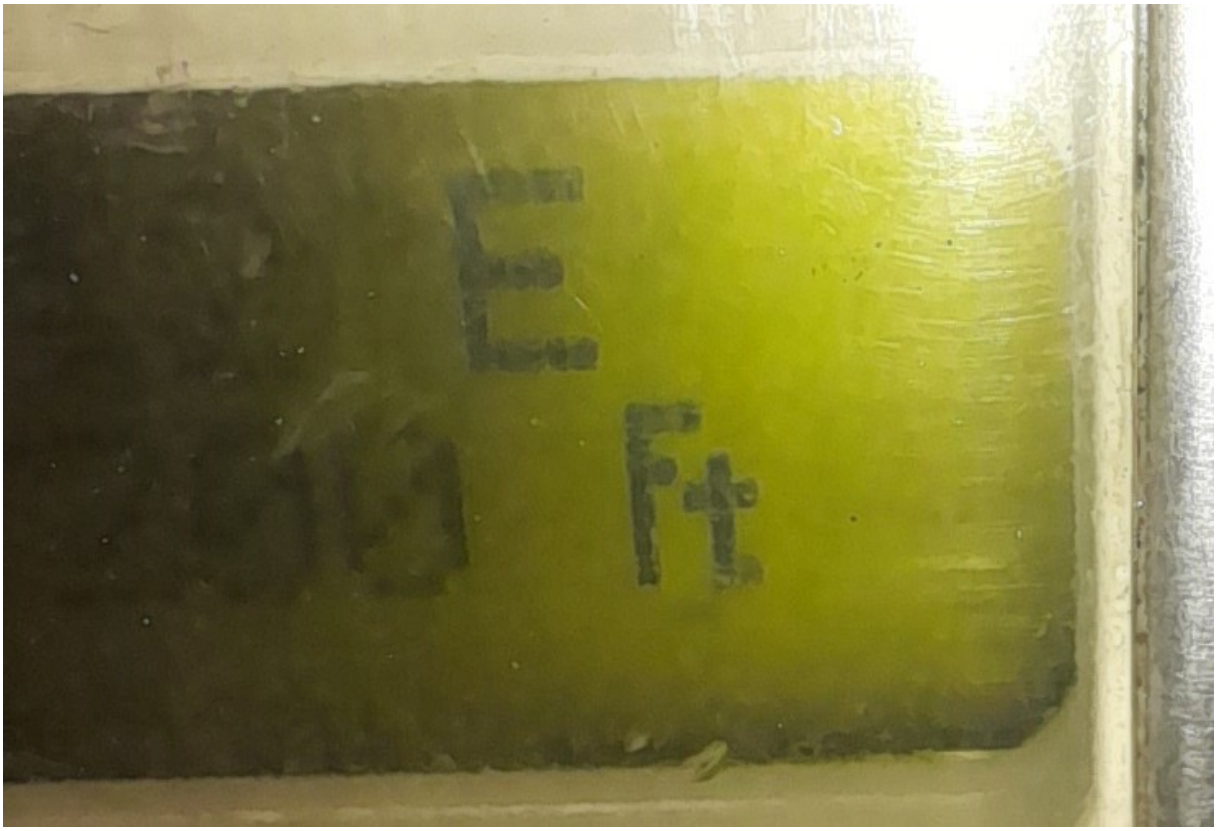

Image 6.1: Close-up of part of the display of a coin-operated public telephone located on Sasadi út, District 11, Budapest, photographed by myself, Vacek Nules, on 23 January 2023. The display shows the denomination of forint coins accepted by the machine, with the largest one in circulation, 200 Ft, listed last. Note Ft sign occupying a single 5×7 character cell.

References

[1] excerpt from *https://www.freedos.org*

[2] see *https://ftp.sun.ac.za/ftp/pub/windows_dos/freedos/files/dos/cpi/cpidos.htm*

[3] see *https://en.wikipedia.org/wiki/CWI-2*

[4] excerpt from *http://primo.homeserver.hu/*

[5] for more about SZTAKI, see *https://www.sztaki.hu/en*

[6] see *http://primo.homeserver.hu/html/ultimoemulator.html*

[7] see *http://primo.homeserver.hu/doc/konyvek/primofuzetek-kezikonyv.pdf*

[8] see *https://www.libri.hu/konyv/TV-Basic-16.html*

[9] see *https://www.youtube.com/playlist?list=PL-OpjJt-qI2sFdfvGG3Ge72akGr1jyplM*

[10] from *https://perkataigyujtemenyek.hu/index.php/gyujtemeny/robotron-irogep*

[11] from *https://www.irogepbolt.hu/images/optima_sp26_irogep.jpg*

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Proposal to Encode a Hungarian Forint Symbol in the Unicode Standard*
2. Requester's name: *Vacek Nules*
3. Requester type (Member body/Liaison/Individual contribution): *Individual Contribution*
4. Submission date: *29-01-2023*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal: *Yes*
   (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): *No*
      Proposed name of script:
   b. The proposal is for addition of character(s) to an existing block: *Yes*
      Name of the existing block: *Currency Symbols*
2. Number of characters in proposal: *1*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary [X] B.1-Specialized (small collection) [ ] B.2-Specialized (large collection) [ ]
   C-Major extinct [ ] D-Attested extinct [ ] E-Minor extinct [ ]
   F-Archaic Hieroglyphic or Ideographic [ ] G-Obscure or questionable usage symbols [ ]
4. Is a repertoire including character names provided? *Yes*
   a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? *Yes*
   b. Are the character shapes attached in a legible form suitable for review? *Yes*
5. Fonts related:
   a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
   Myself, Vacek Nules, if the Committee cannot provide for one
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
   vnules at gmail dot com
6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *Yes*
7. Special encoding issues:
   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*

8. Additional Information:

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/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

[1] Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *No*
   If YES explain
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? *Yes*
   If YES, with whom? *Members of a Hungarian Facebook group for legacy computing enthusiasts*
   If YES, available relevant documents: *Enclosed*
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? *Yes*
   Reference: *Enclosed*
4. The context of use for the proposed characters (type of use; common or rare) *Rare*
   Reference:
5. Are the proposed characters in current use by the user community? *Yes*
   If YES, where? Reference: *On legacy computing platforms which support the character*
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? *Yes*
   If YES, is a rationale provided? *No*
   If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *N/A*
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? *No*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? *No*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? *No*
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? *No*
    If YES, is a rationale for such use provided?
    If YES, reference:
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
    If YES, reference:
12. Does the proposal contain characters with any special properties such as control function or similar semantics? *No*
    If YES, describe in detail (include attachment if necessary)
13. Does the proposal contain any Ideographic compatibility characters? *No*
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: