2023-10-14

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Proposal to disunify Symbols for Legacy Computing from emoji**
Source: **Terminals Working Group**
Authors: **Rebecca Bettencourt, Doug Ewell**
Status: **Individual Contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2023-10-14**

1. Introduction. This document proposes the disunification of 10 graphic characters from legacy computer systems that were previously unified with emoji.

2. History. In April 2017, a list discussion concerning the "PETSCII" character set led to the creation of an ad-hoc Terminals Working Group in charge of proposing characters found in computer systems manufactured in the 1970s and 1980s. The group's first proposal, L2/19-025, was accepted and published in Unicode 13.0. Two subsequent proposals, L2/21-234 and L2/21-235, were also accepted and are scheduled to be published in Unicode 16.0. While writing these proposals, the Terminals Working Group unified as many legacy characters as possible with existing Unicode characters, which happened to include emoji.

In June 2023, Charlotte Buff submitted L2/23-142, a proposal to define variation sequences for 10 characters from legacy computer systems that had been unified with emoji. The proposal was reviewed by the Properties & Algorithms Group, who then forwarded it to the Script Ad-Hoc. The Script Ad-Hoc then contacted two members of the Terminals Working Group, requesting the creation of a proposal to encode these 10 characters separately, which resulted in this document.

3. Rationale. The UTC has determined that unifying text and emoji characters was a mistake, and that attempts will be made to prevent further such unifications. The Emoji Subcommittee even has a standing agreement with the UTC to only propose new characters rather than adding the emoji property to existing characters, and to avoid creating new presentation sequences. The Symbols for Legacy Computing proposals have resulted in the opposite case of needing to use existing emoji characters as non-emoji; regardless, the requirement to prevent emoji unification still applies. Upon reviewing Buff's proposal, the Script Ad-Hoc (including the members most opposed to the Symbols for Legacy Computing proposals) came to the consensus that the best approach is to propose these as separate characters.

4. Proposed disunifications. Proposed new characters are listed on the left. Existing characters are listed on the right. Proposed glyphs have been designed to better match the source characters.

Symbols for Legacy Computing Supplement, gap at U+1CCFA-1CCFF	
U+1CCFA SNAKE SYMBOL	U+1F40D SNAKE
U+1CCFB FLYING SAUCER SYMBOL	U+1F6F8 FLYING SAUCER
U+1CCFC NOSE SYMBOL	U+1F443 NOSE
Symbols for Legacy Computing Supplement, end of range	
U+1CEBA FRAGILE SYMBOL	U+1F377 WINE GLASS
U+1CEBB OFFICE BUILDING SYMBOL	U+1F3E2 OFFICE BUILDING
U+1CEBC TREE SYMBOL	U+1F333 DECIDUOUS TREE
U+1CEBD APPLE SYMBOL	U+1F34E RED APPLE
U+1CEBE CHERRY SYMBOL	U+1F352 CHERRIES
U+1CEBF STRAWBERRY SYMBOL	U+1F353 STRAWBERRY
Symbols for Legacy Computing	
U+1FBFA ALARM BELL SYMBOL	U+1F514 BELL

5. Unicode character properties.

```
1CCFA;SNAKE SYMBOL;So;0;ON;;;;;N;;;;;
1CCFB;FLYING SAUCER SYMBOL;So;0;ON;;;;;N;;;;;
1CCFC;NOSE SYMBOL;So;0;ON;;;;;N;;;;;
1CEBA;FRAGILE SYMBOL;So;0;ON;;;;;N;;;;;
1CEBB;OFFICE BUILDING SYMBOL;So;0;ON;;;;;N;;;;;
1CEBC;TREE SYMBOL;So;0;ON;;;;;N;;;;;
1CEBD;APPLE SYMBOL;So;0;ON;;;;;N;;;;;
1CEBE;CHERRY SYMBOL;So;0;ON;;;;;N;;;;;
1CEBF;STRAWBERRY SYMBOL;So;0;ON;;;;;N;;;;;
1FBFA;ALARM BELL SYMBOL;So;0;ON;;;;;N;;;;;
```

6. References.

Bettencourt, Rebecca et al. 2021. "Proposal to add characters from Smalltalk to the UCS." UTC document L2/21-234. https://www.unicode.org/L2/L2021/21234-terminals-smalltalk.pdf

Bettencourt, Rebecca et al. 2021. "Proposal to add further characters from legacy computers and teletext to the UCS." UTC document L2/21-235. https://www.unicode.org/L2/L2021/21235r-terminals-supplement.pdf

Buff, Charlotte. 2023. "Proposal to Define Variation Sequences for Emoji Mapped to Legacy Computing Symbols." UTC document L2/23-142. https://www.unicode.org/L2/L2023/23142-legacy-computing-emoji-vs.pdf

Ewell, Doug et al. 2019. "Proposal to add characters from legacy computers and teletext to the UCS." UTC document L2/19-025. https://www.unicode.org/L2/L2019/19025-terminals-prop.pdf

1CC0 1CC1 1CC2 1CC3 1CC4 1CC5 1CC6 1CC7 1CC8 1CC9 1CCA 1CCB 1CCC 1CCD 1CCE

0

1

2

3

4

5

6

7

8

9

A

B

C

D

E

F

	1CCF	1CD0	1CD1	1CD2	1CD3	1CD4	1CD5	1CD6	1CD7	1CD8	1CD9	1CDA	1CDB	1CDC	1CDD
0															
1															
2															
3															
4															
5															
6															
7															
8															
9															
A	1CCFA														
B	1CCFB														
C	1CCFC														
D															
E															
F															

	1CDE	1CDF	1CE0	1CE1	1CE2	1CE3	1CE4	1CE5	1CE6	1CE7	1CE8	1CE9	1CEA	1CEB
0														
1														
2														
3														
4														
5														
6														
7														
8														
9														
A														♈ 1CEBA
B														▦ 1CEBB
C														🌳 1CEBC
D														🍎 1CEBD
E														🍒 1CEBE
F														🍓 1CEBF

Terminal graphic characters

Unlike the referenced characters, these characters are not emoji.

1CCFA   🐍   SNAKE SYMBOL
       → 1F40D ☐ snake
1CCFB   🛸   FLYING SAUCER SYMBOL
       → 1F6F8 ☐ flying saucer
1CCFC   👃   NOSE SYMBOL
       → 1F443 ☐ nose

Terminal graphic characters

Unlike the referenced characters, these characters are not emoji.

1CEBA   🍷   FRAGILE SYMBOL
       → 1F377 ☐ wine glass
1CEBB   🏢   OFFICE BUILDING SYMBOL
       → 1F3E2 ☐ office building
1CEBC   🌳   TREE SYMBOL
       → 1F333 ☐ deciduous tree
1CEBD   🍎   APPLE SYMBOL
       → 1F34E ☐ red apple
       → 1F34F ☐ green apple
1CEBE   🍒   CHERRY SYMBOL
       → 1F352 ☐ cherries
1CEBF   🍓   STRAWBERRY SYMBOL
       → 1F353 ☐ strawberry

	1FB0	1FB1	1FB2	1FB3	1FB4	1FB5	1FB6	1FB7	1FB8	1FB9	1FBA	1FBB	1FBC	1FBD	1FBE	1FBF
0																
1																
2																
3																
4																
5																
6																
7																
8																
9																
A																🔔 1FBFA
B																
C																
D																
E																
F																

Terminal graphic characters

Unlike some of the referenced characters, these characters are not emoji.

1FBFA   🔔   ALARM BELL SYMBOL
> → 237E ☐ bell symbol
> → 2407 ᴮᴱᴸ symbol for bell
> → 1F514 ☐ bell

A. Administrative

1. Title
Proposal to disunify Symbols for Legacy Computing from emoji
2. Requester's name
Terminals Working Group (Rebecca Bettencourt et al.)
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2023-10-14
5. Requester's reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
Yes.
6b. More information will be provided later
No.

B. Technical - General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
No.
1b. Proposed name of script
1c. The proposal is for addition of character(s) to an existing block
Yes.
1d. Name of the existing block
Symbols for Legacy Computing Supplement.
2. Number of characters in proposal
10.
3. Proposed category (A-Contemporary; 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)
Category B.1.
4a. Is a repertoire including character names provided?
Yes.
4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes.
4c. Are the character shapes attached in a legible form suitable for review?
Yes.
5a. Who will provide the appropriate computerized font (ordered preference: TrueType, or PostScript format) for publishing the standard?
Rebecca Bettencourt.
5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Rebecca Bettencourt, FontForge.
6a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?
Yes.
6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
7. 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)?
Yes.
8. 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.
See above.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
No.
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

No.
2b. If YES, with whom?
2c. If YES, available relevant documents
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
Contemporary use by specialists and hobbyists.
4a. The context of use for the proposed characters (type of use; common or rare)
Rare.
4b. Reference
5a. Are the proposed characters in current use by the user community?
Yes.
5b. If YES, where?
Worldwide, but particularly in North America, Europe, and Japan.
6a. After giving due considerations to the principles in the P&P document, must the proposed characters be entirely in the BMP?
No.
6b. If YES, is a rationale provided?
6c. If YES, reference
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
Mostly yes, but this is not required.
8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
Yes, all proposed characters could be considered a presentation form of an existing emoji character.
8b. If YES, is a rationale for its inclusion provided?
Yes.
8c. If YES, reference
Included in proposal.
9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
No.
9b. If YES, is a rationale for its inclusion provided?
9c. If YES, reference
10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
Yes.
10b. If YES, is a rationale for its inclusion provided?
Yes.
10c. If YES, reference
Included in proposal.
11a. Does the proposal include use of combining characters and/or the use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1:2000)?
No.
11b. If YES, is a rationale for such use provided?
11c. If YES, reference
11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
11e. If YES, reference
12a. Does the proposal contain characters with any special properties such as control function or similar semantics?
No.
12b. If YES, describe in detail (include attachment if necessary)
13a. Does the proposal contain any Ideographic compatibility character(s)?
No.
13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?