

2009-10-28

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 encode Regional Indicator Symbols in the UCS****Source: Michael Everson and Ken Whistler****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Replaces: Part of N3607, N3680, part of N3690****Date: 2009-09-18**

In document N3607, the German and Irish National Bodies responded to the documents N3582 “Proposal for Encoding Emoji Symbols” and N3583 “Emoji Symbols Proposed for New Encoding” from the Unicode Consortium. The latter proposed to add 10 “Emoji regional indicator symbols” which represent the People’s Republic of China (CN), Germany (DE), Spain (ES), France (FR), the UK (GB), Italy (IT), Japan (JP), Korea (KR), Russia (RU), and the US (US). In N3680, Germany and Ireland proposed encoding 256 characters to represent these and additional countries as defined in ISO 3166. In document N3690, the US National Body reiterated its request for 10 “Emoji compatibility characters”.

The Emoji Ad-Hoc Committee at the WG2 meeting in Tokyo recommended another alternative proposal, which was agreed upon. This document represents that alternative.

The request is to add 26 characters for letters used for representation of regional indicators, which would be used in sequences within, for example, emoji protocols to represent the 10 regional indicators (or possible future extensions). The character names and code positions proposed are:

1F1E6 REGIONAL INDICATOR SYMBOL LETTER A

..

1F1FF REGIONAL INDICATOR SYMBOL LETTER Z.

Character properties proposed are:

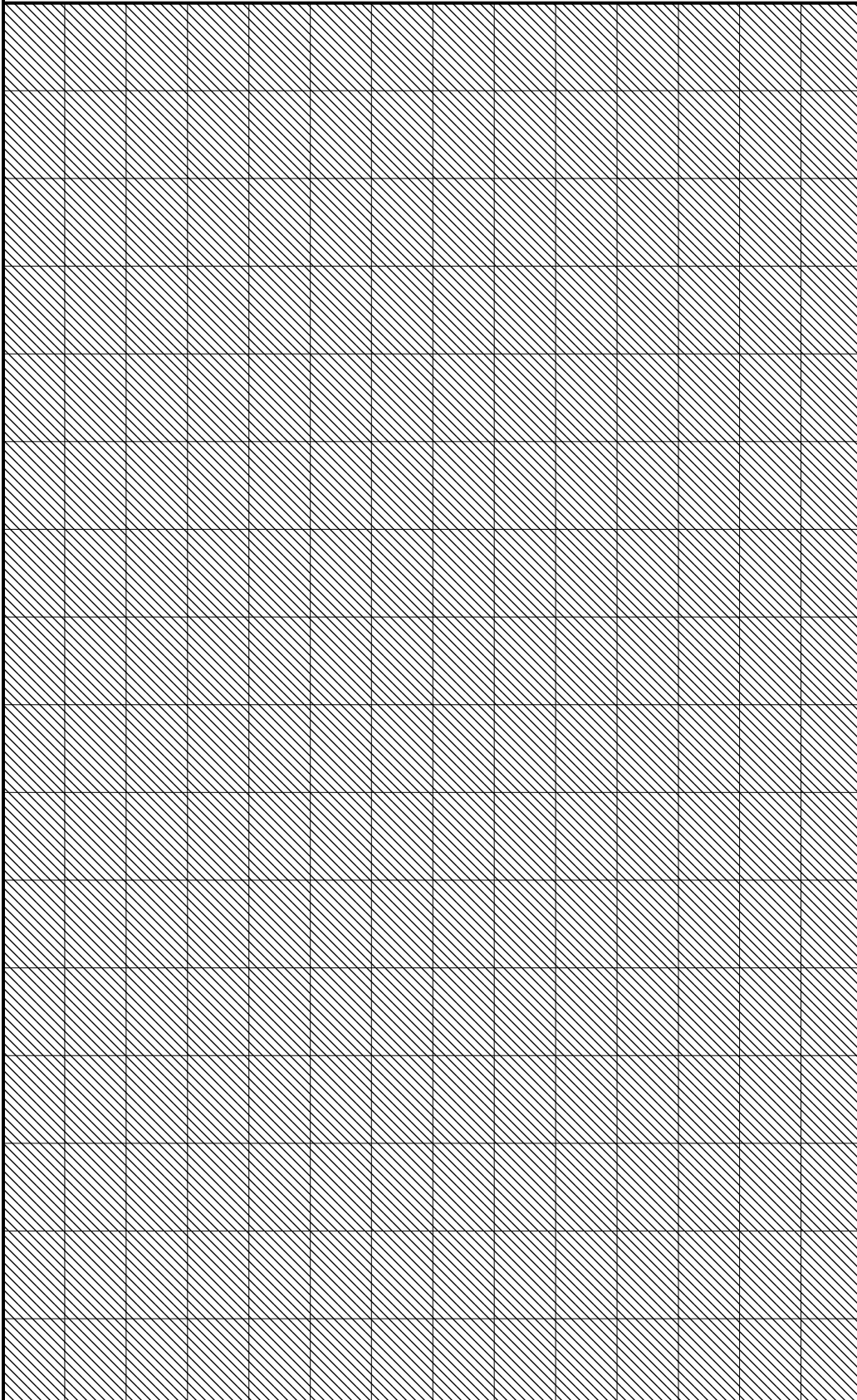
1F1E6;REGIONAL INDICATOR SYMBOL A;So;0;L;;;;N;;;;;

..

1F1FF;REGIONAL INDICATOR SYMBOL Z;So;0;L;;;;N;;;;;

These characters can be used in pairs to represent regional codes. In some emoji implementations, certain pairs may be recognized and displayed by alternate means; for instance, an implementation might recognize F + R and display this combination with a symbol representing the flag of France.



| | 1F10 | 1F11 | 1F12 | 1F13 | 1F14 | 1F15 | 1F16 | 1F17 | 1F18 | 1F19 | 1F1A | 1F1B | 1F1C | 1F1D | 1F1E | 1F1F | | | |
|---|---|------|------|------|------|------|------|------|------|------|------|------|------|------|------|----------|-------|----------|-------|
| 0 |  | | | | | | | | | | | | | | | K | 1F1F0 | | |
| 1 | | | | | | | | | | | | | | | | L | 1F1F1 | | |
| 2 | | | | | | | | | | | | | | | | M | 1F1F2 | | |
| 3 | | | | | | | | | | | | | | | | N | 1F1F3 | | |
| 4 | | | | | | | | | | | | | | | | O | 1F1F4 | | |
| 5 | | | | | | | | | | | | | | | | P | 1F1F5 | | |
| 6 | | | | | | | | | | | | | | | | A | 1F1E6 | Q | 1F1F6 |
| 7 | | | | | | | | | | | | | | | | B | 1F1E7 | R | 1F1F7 |
| 8 | | | | | | | | | | | | | | | | C | 1F1E8 | S | 1F1F8 |
| 9 | | | | | | | | | | | | | | | | D | 1F1E9 | T | 1F1F9 |
| A | | | | | | | | | | | | | | | | E | 1F1EA | U | 1F1FA |
| B | | | | | | | | | | | | | | | | F | 1F1EB | V | 1F1FB |
| C | | | | | | | | | | | | | | | | G | 1F1EC | W | 1F1FC |
| D | | | | | | | | | | | | | | | | H | 1F1ED | X | 1F1FD |
| E | | | | | | | | | | | | | | | | I | 1F1EE | Y | 1F1FE |
| F | | | | | | | | | | | | | | | | J | 1F1EF | Z | 1F1FF |

Regional Indicator Symbols

These characters can be used in pairs to represent regional codes. In some emoji implementations, certain pairs may be recognized and displayed by alternate means; for instance, an implementation might recognize F + R and display this combination with a symbol representing the flag of France.

| | | |
|-------|----------|------------------------------------|
| 1F1E6 | A | REGIONAL INDICATOR SYMBOL LETTER A |
| 1F1E7 | B | REGIONAL INDICATOR SYMBOL LETTER B |
| 1F1E8 | C | REGIONAL INDICATOR SYMBOL LETTER C |
| 1F1E9 | D | REGIONAL INDICATOR SYMBOL LETTER D |
| 1F1EA | E | REGIONAL INDICATOR SYMBOL LETTER E |
| 1F1EB | F | REGIONAL INDICATOR SYMBOL LETTER F |
| 1F1EC | G | REGIONAL INDICATOR SYMBOL LETTER G |
| 1F1ED | H | REGIONAL INDICATOR SYMBOL LETTER H |
| 1F1EE | I | REGIONAL INDICATOR SYMBOL LETTER I |
| 1F1EF | J | REGIONAL INDICATOR SYMBOL LETTER J |
| 1F1F0 | K | REGIONAL INDICATOR SYMBOL LETTER K |
| 1F1F1 | L | REGIONAL INDICATOR SYMBOL LETTER L |
| 1F1F2 | M | REGIONAL INDICATOR SYMBOL LETTER M |
| 1F1F3 | N | REGIONAL INDICATOR SYMBOL LETTER N |
| 1F1F4 | O | REGIONAL INDICATOR SYMBOL LETTER O |
| 1F1F5 | P | REGIONAL INDICATOR SYMBOL LETTER P |
| 1F1F6 | Q | REGIONAL INDICATOR SYMBOL LETTER Q |
| 1F1F7 | R | REGIONAL INDICATOR SYMBOL LETTER R |
| 1F1F8 | S | REGIONAL INDICATOR SYMBOL LETTER S |
| 1F1F9 | T | REGIONAL INDICATOR SYMBOL LETTER T |
| 1F1FA | U | REGIONAL INDICATOR SYMBOL LETTER U |
| 1F1FB | V | REGIONAL INDICATOR SYMBOL LETTER V |
| 1F1FC | W | REGIONAL INDICATOR SYMBOL LETTER W |
| 1F1FD | X | REGIONAL INDICATOR SYMBOL LETTER X |
| 1F1FE | Y | REGIONAL INDICATOR SYMBOL LETTER Y |
| 1F1FF | Z | REGIONAL INDICATOR SYMBOL LETTER Z |

A. Administrative

1. Title

Proposal to encode Regional Indicator Symbols in the UCS

2. Requester's name

Michael Everson and Ken Whistler

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2009-10-28

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

ENclosed Alphanumeric Supplement.

2. Number of characters in proposal

26.

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: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

No.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes, some.

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)?

No.

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. 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/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

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?

General use.

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 in various contexts.

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)?

No.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

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?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or 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?

No.

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?