

**Re: Proposed special case mappings for U+24C2 CIRCLED LATIN CAPITAL LETTER M**  
**From: Unicode Emoji Subcommittee**  
**To: Unicode Technical Committee**  
**Date: 2018-01-18**

U+24C2 is unique among emoji in that it has a case mapping when lowercased. Since emoji-variation-sequences.txt defines the usual pair of variation sequences for text and emoji presentation, we have the following results:

toLowerCase(...)	result
Ⓜ <24C2>	Ⓜ <24DC>
Ⓜ + text VS <24C2 FE0E>	Ⓜ + text VS <24DC FE0E>
Ⓜ + emoji VS <24C2 FE0F>	Ⓜ + emoji VS <24DC FE0F>

Since casing operations are not affected by variation selectors, the results include two invalid variation sequences (in red), and in the latter emoji presentation has been lost.

### Solution

We propose the following VS-sensitive-case-mappings (in yellow):

toLowerCase(...)	result
Ⓜ <24C2>	Ⓜ <24DC>
Ⓜ + text VS <24C2 FE0E>	Ⓜ <24DC>
Ⓜ + emoji VS <24C2 FE0F>	Ⓜ + emoji VS <24C2 FE0F>

The rationale for these mappings is that they preserve the presentation (text or emoji) of the original sequence and avoid invalid variation sequences. Preventing the lowercasing of U+24C2 with emoji presentation should be unsurprising to users, since no other emoji containing Latin letters (e.g. 🇦🇪, 🇦🇰) have case mappings.

### Implementation

a) In <https://www.unicode.org/Public/UCD/latest/ucd/SpecialCasing.txt>, if the first entry can be a sequence, then add the following lines:

```
# emoji characters with variation selectors
 24C2 FE0E; 24DC; 24C2 FE0E; 24C2 FE0E; # CIRCLED LATIN CAPITAL LETTER M
+ VARIATION SELECTOR-15 (text style)
 24C2 FE0F; 24C2 FE0F; 24C2 FE0F; 24C2 FE0F; # CIRCLED LATIN CAPITAL
LETTER M + VARIATION SELECTOR-16 (emoji style)
```

b) Otherwise, we might need to add a context type for special casing; see:

- 3.13 Default Case Algorithms, p 153 ff in

<http://www.unicode.org/versions/Unicode10.0.0/ch03.pdf>

- 5.18 Case Mappings, p 238 ff in <https://www.unicode.org/versions/Unicode10.0.0/ch05.pdf>