

Feedback regarding Unicode 18 additions

Eduardo Marin Silva

02/06/2026

I understand that much of this feedback is coming late but I have been very busy with work. In any event, I have left my pieces of feedback for what they are worth.

1) Sheva Na: The [original proposal](#) indicates there should be an informative alias “= mobile sheva”

2) Compound Tone diacritics: I would add the following reciprocal cross references:

New addition	Cross references
1AEE COMBINING DOUBLE GRAVE ACCENT BELOW	030F COMBINING DOUBLE GRAVE ACCENT, 0316 COMBINING GRAVE ACCENT BELOW
1AEF COMBINING DOUBLE ACUTE ACCENT BELOW	030B COMBINING DOUBLE ACUTE ACCENT, 0317 COMBINING ACUTE ACCENT BELOW

I would also like to mention that the cross reference under 1AF0 (double comma above), to 1AB8 (double open mark below), seems like a stretch and can be removed.

3) Currency Signs: Consider adding the note under 20C3 stating “• United Arab Emirates”, as not everybody would be able to tell at first what “UAE” stands for. I would also add reciprocal cross references under 20C1 (Saudi Rial) to FDFC (Legacy Rial).

4) Wiggly Exclamation Mark: The note under 2E60 stating “• wiggles are broader at top than at bottom” is not useful, as it’s readily apparent from the representative glyph. It’s clear that having a punctuation mark with a similar glyph, but with consistent width for the wiggles would be a good candidate for unification with this character. Consider removing the note.

5) Capital Closed Omega: Character A7DD is currently under the header “Letters used in Wakashan and Salishan languages” which is clearly incorrect. Add a new header for it “Letter for the English Phonotypic Alphabet”.

6) Arabic Small Circle Above: Character 10ECE (Arabic small circle above) was disunified from 0652 (Sukun Above) but back in version 14, a similar situation occurred, with a pair of combining different sized circles (08D0 & 08D1), with the difference being they were placed below. The character for the smaller circle was unified with a positional variant of the sukun (hence the name of 08D0 being ARABIC SUKUN BELOW). The small circle above can be unified with the regular sukun, unless there is a technical reason not to.

7) Devanagari Letter Alternate Ddda: I propose a clearer wording for the note under this character “• alternate form of 097E preferred in some old texts, but not in current use”; along with reciprocal cross references to 097E.

8) Bengali characters:

- a) Add reciprocal cross references between the new character 11DF0 (Bengali Anusvara Above) and the regular one 0982 (Spacing Bengali Anusvara).
- b) Add reciprocal cross references between 11DF1 (Sanskrit Ba) and 09F0 (Assamese Ra).
- c) Add a note under 0982 stating “• a glyph variant without the diagonal stroke was used in old manuscripts” as per [this document](#).

9) Khitan Small Script Additions: Under 18C3E I would replace the current notes with the following: “• the only attestation of this character in a primary source, is the same *paizi* that is tentatively identified as being written in the Jurchen Small Script” & “• the glyph reflects the appearance in the primary source”.

Although the second note can be omitted, as the glyph difference is negligible; but if it’s considered significant enough to merit a note, then there should be blue highlighting.

10) Historic Katakana: There should be a cross reference from 1B128 (Old Katakana Wi) to 1B10D (Hentaigana Wi).

11) Radix signs: The names of these characters while technically correct are not ideal. In the sense that if ones wants to tell these signs apart by name, the user needs to count the instances of the repetition of the term “square root” which is not easy to do when it comes to pieces of text rather than simple shapes. I propose the following:

Current name	Proposed name
1CEDD SQUARE ROOT OF SQUARE ROOT	1CEDD SQUARE ROOT WITH HUMP
1CEDE SQUARE ROOT OF SQUARE ROOT OF SQUARE ROOT	1CEDE SQUARE ROOT WITH DOUBLE HUMP
1CEDF SQUARE ROOT OF SQUARE ROOT OF SQUARE ROOT OF SQUARE ROOT	1CEDf SQUARE ROOT WITH TRIPLE HUMP

I would also add reciprocal cross references from 1CEFC (angle opening up) to 2220 (baseline angle).

12) Musical symbols: For these there are several suggestions

- a)** Replace the cross reference under 1D127 from 0301 (acute accent above) to 0317 (acute accent below). Similarly, replace the cross reference under 12128 from 0306 (breve above) to 032E (breve below).
- b)** Above the character 1D1E9 there are two headers unnecessarily. Consolidate both into one header that says “Iranian quarter tone accidentals” and remove the note under the first header.
- c)** Move the header that says “Dynamics” from above 1D18F to above 1D18C
- d)** Add reciprocal cross references for the following pairs:

1D101 MUSICAL SYMBOL DOUBLE BARLINE	1D262 MUSICAL SYMBOL HEAVY DOUBLE BARLINE
1D110 MUSICAL SYMBOL FERMATA	1D26B MUSICAL SYMBOL TRIANGULAR FERMATA
1D112 MUSICAL SYMBOL BREATH MARK	1D272 MUSICAL SYMBOL SALZEDO BREATH MARK
1D11E MUSICAL SYMBOL G CLEF	1D276 MUSICAL SYMBOL DOUBLE G CLEF
1D13A MUSICAL SYMBOL MULTI REST	1D1FF MUSICAL SYMBOL LONGA REST
1D183 MUSICAL SYMBOL ARPEGGIATO UP	1D27F MUSICAL SYMBOL ARPEGGIATO UP WITH ARROW
1D172 MUSICAL SYMBOL COMBINING FLAG-5	1D250 MUSICAL SYMBOL COMBINING FLAG-6
1D142 MUSICAL SYMBOL ONE HUNDRED TWENTYEIGHTH REST	1D253 MUSICAL SYMBOL TWO HUNDRED FIFTY-SIXTH REST
1D18F MUSICAL SYMBOL PIANO	1D264 MUSICAL SYMBOL NIENTE
1D1B9 MUSICAL SYMBOL SEMIBREVIS WHITE	1D268 MUSICAL SYMBOL DIAMOND NOTEHEAD WHITE
1D1BA MUSICAL SYMBOL SEMIBREVIS BLACK	1D269 MUSICAL SYMBOL DIAMOND NOTEHEAD BLACK

- d)** The glyph of 1D25F (buzz roll stem) should show the stem positioned above right of the dotted circle to be consistent with similar characters

13) Leibnizian operators:

- a)** Add reciprocal cross references from 1DB00 (square equals sign) to 2293 (square cap), another one from 1DB1A (similarity sign) to 3030 (wavy dash) and another one from 1DB1C (facit symbol) to 2A0D (finite part integral)
- b)** Why was 1DB1B (dissimilarity sign) not considered a variant of 2241 (negated tilde)? In any event they may benefit from a cross reference.

c) Is 1DB10 (Cartesian Equals) meant to be taken as the reversed version of 221D (Proportional To)? In which case they need reciprocal cross references and making sure the BIDI file is updated.

14) Miscellaneous Latin letters:

a) **1DF57 LATIN SMALL LETTER TURNED R WITH MIDHEIGHT LEFT HOOK** does not have an accurate name, as in the inverted position, the hook is no longer “mid-height”. Consider instead “LATIN SMALL LETTER TURNED R WITH HOOK ABOVE BASELINE”; name was chosen to avoid clashing with 027B LATIN SMALL LETTER TURNED R WITH HOOK and so they may benefit from reciprocal cross references. The change of name would also affect the modifier variant located at 1DFCD.

b) **1DF60 LATIN SMALL LETTER OMEGA WITH LOOP** has a cross reference to A74F (Double O) for some reason. Regardless of if that reference is merited or not, I would add a reciprocal cross reference to A7B7 (the Regular Latin Omega)

c) **1DF6F LATIN SMALL LETTER PHONOTYPIC DIPHTHONG AI** can have cross references to 0268 (I with stroke) and 1D96 (I with Retroflex Hook)

d) **Non Casing Capital Letters:** Some capital letters had to be encoded without a corresponding lowercase, but the are informal lowercase mappings must be mentioned at least: 1DF80 (Capital A with Topbar) should have a note “• the informal lowercase is considered 024B (Small Q with Hook)”. For 1DF81 (Capital E with Bent Topbar) there should be a note stating “• the informal lowercase is considered 1EB9 (Small E with Dot Below)”. The header should

e) **1DF95 (LATIN SMALL LIGATURE LONG S WITH DESCENDER S:** The name of this character is not very clear; naming it LATIN SMALL LETTER SHARP S WITH DESCENDER is better. Also I noticed that they are also proposing a different codepoint for an italic version of the character (1D6A6); however that would only be merited if the regular and the italic versions were used contrastively in mathematical contexts, but the original [proposal document](#) makes no such claim.

15) Geometric shapes: Characters 1F7FD & 1F7FE were disunified from 1CC86 & 1CC87 respectively. After some thought, I speculate that maybe it’s because while the former must have the base of the triangles sitting on the baseline, the latter must have it sitting at the bottom of the character cell. I’m not sure, but if that’s the case then they could benefit from reciprocal cross references.

16) New emoji:

a) **Monarch Butterfly:** The argument for encoding this character atomically is not very strong in my opinion, because it can be made sequence 🦋 1F451 + ZWJ + 🦋 1F98B. This is fine as it would fall back to “🦋🦋” which gets the meaning across. This would force vendors to use a generic glyph for 1F98B. This is consistent with the decision from Unicode to encode a “Phoenix” with the sequence 1F426 🦋 + ZWJ + 1F525 🔥

b) **Pickle:** Why is the disunification with the cucumber 1F952 🥒 warranted? A pickle is just a cucumber prepared a certain way, and they lack visual distinctiveness. If Unicode feels like unifying the squirrel and the chipmunk is warranted why not use the same codepoint for both concepts? They could add CLDR alias to 1F952 and even add a note stating that it can represent pickles as well.