

Anomalies regarding the representative glyphs of some characters

Eduardo Marin Silva

26/10/2025

While working on a separate document regarding the code charts I discovered that some representative glyphs have an unexpected appearance and I would like to bring attention to that. First I describe the anomalies and then I propose a solution.

1. **Arrowhead modifier letters:** 02C2-02C5 Contains letters in the shape of arrowheads. Wikitionary claims they are used for NAPA, but I couldn't verify this myself. The representative glyphs are large, even confusable to characters like less than 003C < or greater than 003E >.

02C2	<	MODIFIER LETTER LEFT ARROWHEAD
		• fronted articulation
02C3	>	MODIFIER LETTER RIGHT ARROWHEAD
		• backed articulation
02C4	^	MODIFIER LETTER UP ARROWHEAD
		• raised articulation
		→ 005E ^ circumflex accent
		→ 2303 ^ up arrowhead
02C5	∨	MODIFIER LETTER DOWN ARROWHEAD
		• lowered articulation

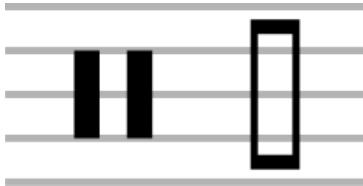
The original submission documents are either lost or very hard to read, as I cannot find the original justification for this presentation or captures of the original attestations. If they are meant to be this large, then cross references to similar characters are warranted, but if they are meant to be superscript then the cross references that need to be added changes. Most fonts use the representative glyphs of Unicode but some fonts like the Noto family (^~) or DejaVu (^~) depict the characters as superscript which complements nicely with the "low" arrowhead letters (02EF-02F2) and the combining left and right arrowheads above 1DFE – 0350 –. The superscript presentation does introduce confusability with 02C6 ^ & 02C7 ^ but it's evident that the characters are typographically distinct regardless, and the glyphs can be made to accentuate the difference

2. **Arrowhead symbols:** One would expect symbols 2303 ^ & 2304 ∨ to simply be rotated versions of each other, but for some reason the representative glyph of 2303 is noticeable wider than 2304.

2303 ^ UP ARROWHEAD

2304 ∨ DOWN ARROWHEAD

3. **Low inverted exclamation mark:** A71F i is a graphic alternate of A71E i to represent tonal upstep in Africanist texts. It's intended for use when other characters like modifier small i (i), would be confusable, However, despite the original attestations in the proposal document ([N2945](#)) showing a clearly subscript glyph (i), the representative glyph has it on the midline.
4. **Drum Clefs:** 1D125 || & 1D126 || are meant to represent a drum clefs but if one is to google "drum clef" the images that appear are very different glyphs.



The current representative glyphs are confusable with other musical symbols: 1D101 || (double barline) & 1D15A □ (white note head cluster).

1D125		MUSICAL SYMBOL DRUM CLEF-1
1D126		MUSICAL SYMBOL DRUM CLEF-2

5. **Small g variants:** In Unicode there is a distinction between the regular small letter g (0067) that may have an open or closed tail depending on the style (g or ḡ) and the 'script g' (0261) that is always 'open' regardless of the style (ḡ). The letters are further contrasted by their different uppercase forms (G vs Ḡ). This in itself is not problematic, but since the representative glyph for the 'regular g' is with a closed tail, then one would expect all variants of g to follow that style to maintain consistency, unless the name specifies otherwise (this would exclude AB36 ḡ). But two of the g variants use the open style despite the name being generic.

0260	ḡ	LATIN SMALL LETTER G WITH HOOK
		<ul style="list-style-type: none"> • implosive velar stop • uppercase is 0193 Ḡ

1D83 ḡ LATIN SMALL LETTER G WITH PALATAL HOOK

6. **Urdu Arabic letters:** In the code charts of any cursive joining script (including Arabic), it's almost universal that the representative glyphs shows the character in isolated form. There are some exceptions, for instance if there is no proper isolated form attested, (like the Arabic Vertical Tail: 088E ݚ), the character is a presentation form encoded for legacy compatibility (these would include most of the characters on both the Arabic Presentation Forms blocks), or the character is a presentation form, disunified and encoded atomically for technical simplicity, (like the Old Uyghur Final Heth: 10F75 ݚ). However the Arabic letters 06C1 (heh goal), 06C2 (heh goal with hamza above) & 06C3 (teh marbuta goal) have their final form as the representative glyphs despite having different looking isolated forms.

06C1	~	ARABIC LETTER HEH GOAL
		<ul style="list-style-type: none"> • Urdu
06C2	ؚ	ARABIC LETTER HEH GOAL WITH HAMZA ABOVE
		<ul style="list-style-type: none"> • Urdu • actually a ligature, not an independent letter ≡ 06C1 ~ 0654 ۅ
06C3	ؚ	ARABIC LETTER TEH MARBUTA GOAL
		<ul style="list-style-type: none"> • Urdu

7. **Double stacked comma:** This character (2E49 ;) was added for transcribing Church Slavonic liturgical texts. The representative glyph therefore matches the appearance the character has on those texts.

2E49 ; DOUBLE STACKED COMMA

However, it's implied that the regular comma (002C ,) is unified with the comma used on those texts, even when the appearance is more like 2E49. Should we find two commas on top of each other with the 9-like appearance in other contexts (;) would that be a different character that needs to be disunified?

Proposed solutions. I shall now elaborate on what I believe to be the most efficient resolutions, from most simple to most involved.

- For number 2, it's clear it was just a mistake by the editors. I believe that 2303 should have its glyph compressed. No glyph erratum notice would be necessary, as the fonts I tested seems to keep the glyphs consistent regardless.
- For number 3, it's clear it was just a mistake by the editors. This mistake however has resulted in many fonts positioning the glyph incorrectly. This is important, as the entire point of the alternate character was to increase visual distinctiveness, but if it's just lowered halfway, the effect is reduced. I recommend the glyph be corrected along with a glyph erratum notice.
- For number 7, while its understandable why this particular presentation was chosen, using a shape that matches the representative glyph of 002C keeps it explicit that the Typicon forms are unifiable.
- For number 4, it's unclear why the current glyphs were chosen by the original submitter as the document makes no notice of that, but it's clear that they are not the most common. While the differences of the glyph for 1D126 are subtle, they are quite stark with 1D125, and I'm surprised nobody else has taken an issue with it before.
I recommend to change the glyph of 1D125 to be more similar to the image (shorter, wider and with thicker more separated bars) along side a cross reference to 23F8 ॥.
1D126 would just need to be made slightly shorter and the horizontal strokes need to be slightly thicker than the vertical ones. Many fonts have designed these characters with incorrect glyphs because of this, so a glyph erratum notice is warranted for at least 1D125.
- Number 6 seems to be a conscious decision, as 06D5 ܀ ARABIC LETTER AE has a confusable glyph in isolated presentation to 06C1 ܁. However, this is inconsistent with other characters that are also confusable in isolated presentation, like 0649 ܂ ALEF MASKURA & 06CC ܂ FARSI YEH or 066F ܂ DOTLESS QAF & 08BC ܂ AFRICAN QAF.
I recommend changing the glyphs to the isolated presentation. The current glyphs imply that the characters are only attested in final presentation (like 088E) which is incorrect.
While no glyph erratum is necessary, it should be documented somewhere in the release notes. Further confusion can be avoided by better notating the differences between 06C0 & 06C2 and also 06D1 & 06D5.
- For number 5, I withhold judgment, as this would need an inspection of the original sources of these characters and consulting the user community, to decide on the best glyph. I just point this out so that the experts are aware and it's documented. If it's

decided the glyphs need to be changed for consistency, then I recommend also issuing a glyph erratum notice.

- For number 1, It's difficult to come up with a recommendation without looking at the original materials. I hope that this document can inspire the experts to look into this issue. If it turns out the glyphs are correct after all, I have no issue. But I would like confirmation so I can be certain of my cross reference suggestions. Otherwise they glyphs need to be corrected along with a glyph erratum notice