Title: Rationale for commenting out the variation sequence **U+1333B U+FE00** for Ancient Egyptian
From: Peter Dils, Andrew Glass, Jorke Grotenhuis, Svenja Gülden, Mark-Jan Nederhof, Stéphane Polis, Daniel A. Werning
**Date:** 2024-06-15

### Abstract

A 90° rotation of the hieroglyph ⟨glyph⟩, which was introduced earlier as a variation sequence, is found to be inconsistent within the broader context of uniformly encoding rotations of the base sign ⟨glyph⟩. Hence it is best commented out.

Background

- Andrew Glass et al. (2021), *Additional control characters for Ancient Egyptian hieroglyphic texts*, L2/21-248.

- Daniel A. Werning (2022), *Rotations of Egyptian Hieroglyphs to be Registered in Unicode*, L2/22-012R.

- Mark-Jan Nederhof (2024), *Rotations for Ancient Egyptian – a critical analysis*, L2/24-045.

- Daniel A. Werning et al. (2024), *Additional variation selectors for rotations of Ancient Egyptian hieroglyphic texts*, proposal currently under consideration.

- Mark-Jan Nederhof et al. (2024), *A 3-tier system for Ancient Egyptian hieroglyphs*, proposal currently under consideration.

Rotations

as individual code points

Some hieroglyphs occasionally occur in a rotated form. Until recently, rotated forms could only be encoded using individual code points. One sign that is rotated very frequently is the 'hoe' ⟨glyph⟩ = **U+13338**. It is widely known as **U6** from the sign list in Alan Gardiner's *Egyptian Grammar* (1957). In its most common orientation, the top segment appears roughly diagonal, but there is a continuum of possible angles from horizontal to almost vertical. Horizontal forms are relatively common, as are almost vertical forms. The horizontal form ⟨glyph⟩ is widely known as **U7**, which Gardiner listed as 'alternative form of last [= **U6**]', and which exists as individual code point **U+1333B**. The almost vertical form ⟨glyph⟩ exists as individual code point **U+13339**. Its mirror image ⟨glyph⟩ exists as **U+1333A**. However, ⟨glyph⟩ and ⟨glyph⟩ were not listed explicitly in the *Egyptian Grammar*, but appeared in Gardiner's *Supplement to the catalogue of the Egyptian hieroglyphic printing type* (1953). Note that widely used hieroglyphic typesetting tools like JSesh[1] have primitives to rotate signs by arbitrary angles and therefore never needed to list ⟨glyph⟩ and ⟨glyph⟩ among the available signs.

The reader is here reminded that hieroglyphic text consists of groupings of signs, arranged above or below one another, or arranged in a number of other ways. This implies that ⟨glyph⟩, which is low and wide, is not interchangeable with ⟨glyph⟩, which is narrow and tall. For example, if an occurrence of the hoe appears underneath other tall signs, then only ⟨glyph⟩ might reasonably fit within the available space. It is therefore important to be able to encode rotation. However, as there is a continuum of possible angles that may

---

[1] http://jseshdoc.qenherkhopeshef.org

appear in actual inscriptions, this leads to the question to what level of precision one would wish to express angles of rotation. As argued by Nederhof (2024), it is not appropriate for Unicode to express overly small differences in angle, and for any given sign, at most 8 (approximate) angles of rotation between 0° and 360° will be distinguished.

The perfectly vertical form $\not{A}$ is uncommon and never existed as individual code point. See Appendix A for a selection of occurrences of the hoe found in two volumes of line drawings.

as variation sequences

The concept of encoding rotations using variation selectors was introduced by Glass et al. (2021). A list of rotations found in JSesh corpora was compiled by Werning (2022). In a first phase, a number of variation sequences were introduced that correspond to rotations by 90°, 180° or 270°. This included the 90° rotation of ⌐ to give $\not{A}$ ← **U+1333B U+FE00**. The motivation was that rotations of ⌐ of around 90° occurred frequently in JSesh corpora; as noted earlier, $\not{A}$ did not exist as individual sign in JSesh. However, this does not imply that the perfectly vertical form $\not{A}$ is common in original inscriptions, and it appears that encoders in the past chose a 90° rotation of ⌐ as the easiest way to get a more or less satisfactory appearance for what might in reality be an almost vertical orientation $\not{A}$.

Recently, the remaining rotations from Werning (2022) were revised, leading to the proposal by Werning et al. (2024). This includes rotations by angles that are not multiples of 90°. At the same time, a consensus was reached that it is preferable for a rotated form to be henceforth encoded as the combination of a fixed *base* sign and a variation selector, even if that rotated form already existed as an individual code point. This benefits statistical analysis of corpora, it makes text search easier, and text input systems become more uniform.

The base sign should be chosen to be the most common and recognizable form. In this way, users will know, for example, what sign to enter in a search query. Other forms that existed as individual code points will become 'legacy', as proposed by Nederhof et al. (2024), that is, these code points will still be implemented in fonts, but they should be systematically avoided in specialist corpora.

Conclusions

The only appropriate base sign for the hoe is $\not{A}$ = **U+13338**. Other orientations should henceforth be obtained using variation selectors. The other individual code points become legacy, including ⌐ = **U+1333B**. Allowing **U+1333B U+FE00** to coexist with **U+13338 U+FE03** would be undesirable, for several reasons. First, this would mean there is no (single) base form for the hoe, which would nullify the benefits of having normalized encodings with unique base forms. Second, the two forms are very similar, one being almost vertical and the other being perfectly vertical, a distinction that is too small for Unicode to reasonably make. After commenting out **U+1333B U+FE00**, we obtain these encodings:

orientation	present	future	
$\not{A}$	U+13338	$\not{A}$	= U+13338
$\not{A}$	U+13339	$\not{A}$ $\boxed{\overline{\text{VS}}\atop 4}$	= U+13338 U+FE03
$\not{A}$	U+1333A	$\not{A}$ $\boxed{\overline{\text{VS}}\atop 4}$ $\boxed{\leftrightarrow}$	= U+13338 U+FE03 U+13440
⌐	U+1333B	$\not{A}$ $\boxed{\overline{\text{VS}}\atop 7}$	= U+13338 U+FE06
$\not{A}$	U+1333B U+FE00	replace by $\not{A}$	= U+13338 U+FE03

A Selected occurrences in line drawings of stelae

In right-to-left text:

[1, Plate 15] [2, Plate 37] [2, Plate 7] [2, Plate 9] [2, Plate 13]

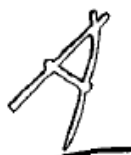[2, Plate 13] [2, Plate 13] [2, Plate 18] [2, Plate 35] [2, Plate 40] [2, Plate 5]

[2, Plate 7] [2, Plate 8] [2, Plate 40] [1, Plate 17]

In left-to-right text:

[1, Plate 45] [2, Plate 35] [2, Plate 13] [2, Plate 20] [2, Plate 35] [2, Plate 5]

[2, Plate 8] [1, Plate 18] [1, Plate 17]

References

[1] M.L. Bierbrier. *Hieroglyphic texts from Egyptian stelae, etc., Part 10.* British Museum, 1982.

[2] T.G.H. James. *Hieroglyphic texts from Egyptian stelae, etc., Part I.* British Museum, 2nd edition, 1961.