

Changes to the text of TUS Core Spec 11.4.2 Egyptian Hieroglyph Format Controls relating to the ordering of elements joined by the overlay control (U+13436)

Andrew Glass — July 30, 2026

During UTC #188 we held an ad-hoc discussion to revise the text of 11.4.2. Participants were (alphabetical order): Andrew Glass, Asmus Freytag, Debbie Anderson, Ken Whistler, Michel Suignard, Robin Leroy, Roozbeh Pournader. We agreed on the following new content to improve the current text:

11.4.2 Egyptian Hieroglyph Format Controls

<https://core-spec.unicode.workers.dev/chapter-11/#G32499>

Hieroglyphs may also overlay other hieroglyphs. This arrangement is controlled by U+13436 ☐ EGYPTIAN HIEROGLYPH OVERLAY MIDDLE. This control character causes a following hieroglyph to overlay on top of a preceding hieroglyph, as shown in the fifth example in Figure 11-2. Hieroglyphs that overlay one another stack at their center points. **Overlay sequences follow an ordering rule which may sequence the component characters differently to their reading order: when a sequence includes two characters that are joined by the overlay control (U+13436 ☐ EGYPTIAN HIEROGLYPH OVERLAY MIDDLE), they are ordered by their kEH_Cat values in ascending order. This rule applies to sequences involving single characters as components of an overlay. It also applies to groups of characters that participate as elements in an overlay.**

For example, the sign ⲙ (U+13138 EGYPTIAN HIEROGLYPH F050) has an alternate encoding as the sequence ☐ ☐ <U+13132, U+13436, U+132F4>. In this sequence, U+13132 ☐ EGYPTIAN HIEROGLYPH F046 comes first, since it has kEH_Cat=F-28-004 and the character U+132F4 ☐ EGYPTIAN HIEROGLYPH S029 comes at the end, since it has kEH_Cat=S-15-019.

In the case of overlaid signs where one overlaid element is made of a group of signs, the same ordering principle applies. For example, the sign ⲙ (U+1325F) may also be represented as the sequence ☐ ☐ ☐ ☐ ☐ ☐. The elements participating in the overlay are and their kEH_Cat values are:

1.  (U+13257 EGYPTIAN HIEROGLYPH O006), kEH_Cat = **O-03-005**
2.      <U+13437 EGYPTIAN HIEROGLYPH BEGIN SEGMENT, U+133CF EGYPTIAN HIEROGLYPH X001, U+13430 EGYPTIAN HIEROGLYPH VERTICAL JOINER, U+1327B EGYPTIAN HIEROGLYPH O029, U+13438 EGYPTIAN HIEROGLYPH END SEGMENT> kEH_Cat = **X-01-001** and O-16-024.

Thus element (1) comes first because **O-03-005** precedes the group (2) which begins with **X-01-001**. In the event of a tie, subsequent members of the group are considered. If all signs in a group are identical, the code point of the joiners should break the tie.

Different authors may analyze complex quadrats in distinct ways, particularly in connection with the overlay middle control. Beyond the ordering requirements detailed above, nothing prevents one from managing these distinctions at the level of orthographic rules.

See also the example of ‘b in Complex Clusters below.

Subheading: Complex clusters

<https://core-spec.unicode.workers.dev/chapter-11/#G32726>

For example, U+13217  EGYPTIAN HIEROGLYPH N035A looks like a stack of three copies of U+13216  EGYPTIAN HIEROGLYPH N035 and could be represented by the sequence <13216, 13430, 13216, 13430, 13216>. However, this compound sign is a logograph for the word for water, mw, whereas the parts are phonemograms with the unrelated value n. As a result, the atomic character U+13217 is preferred. In contrast, consider U+130C1  EGYPTIAN HIEROGLYPH D059, which looks like U+1309D  EGYPTIAN HIEROGLYPH D036 (kEH_Cat D-23-001) over U+130C0  EGYPTIAN HIEROGLYPH D058 (kEH_Cat D-35-012), so that it can be represented as the sequence <130C0, 13436, 1309D>. U+130C1  is a phonemogram with the value ‘b, and the parts are phonemograms whose value make up ‘b — U+1309D  has the value ‘ and U+130C0  has the value b. In this case, **the sequence <1309D, 13436, 130C0> is preferred**. For information on the function and value of an individual hieroglyph, as well as descriptions of complex hieroglyphs in terms of atomic parts, see Unicode Standard Annex.