The Universal Hieroglyphic Writing System: Consensus and possible compromise

DRAFT for Feedback. Bob Richmond 2016-08-03

Mark-Jan, Stéphane and I will be participating remotely in a UTC meeting on Thursday (late afternoon European time).

I've attempted to summarise the current situation on where we are on consensus issues.

I've stated what it seems is the only compromise solution available at this time.

Mark-Jan and Stéphane have not been available this week so I may not have represented their opinions as they would wish – I'll update this if I hear from them before the meeting.

Consensus Status

1. L2/16-018 with EGYPTIAN HIEROGLYPH HORIZONAL JOINER and EGYPTIAN HIEROGLYPH VERTICAL JOINER is an acceptable basis for hieroglyphic writing in plain text.

TLA and Ramses project were comfortable with this in their April letters to UTC. This remains a consensus item.

Objections so far: None.

2. Concerns still exist about EGYPTIAN HIEROGLYPH LIGATURE JOINER

Summary. TLA and Ramses projects along with Nederhof et al stated they had general concerns in their April 2016 letters to UTC. The vast majority of typeset works in print use horizontal text and can be encoded using LIGATURE along with the other two. Only a small proportion of quadrats use LIGATURE (< 5%). No detailed list of quadrats for which LIGATURE is inadequate has been submitted by Egyptologists or others. Some would prefer to use a geometrical approach of one form or another instead but no proven approach has

been suggested yet. Standard monogram-like ligatures such as and could be added to repertoire instead of using controls at all. Some rare complex ligatures have been identified but it is so far unclear whether it is necessary for Unicode to support them all. There are objections to having one method for representing combinations that are a clear part of the writing system and another for rare cases because that would lead to more than one sequence representing the same quadrat. This needs to be resolved on the basis of solid data and investigation. LIGATURE may need to be supplemented or replaced to accomplish what is seen as needed.

Comments/additions/corrections?

3. Additional control characters relating to repertoire extension issues

Control characters such as stack/monogram joiner/insert centre have been floated as alternatives to encoding certain hieroglyph combinations as distinct characters (as is currently the policy for repertoire extension). These could be added as controls at a later stage but are not essential for an initial scheme.

Comments/additions/corrections?

4. Vertical writing and 'tall group' support need full attention

The 3 control solution focuses on horizontal text. This is typically what scholars and others have used for the vast majority of work to date. There is a view Unicode gives an opportunity to encourage new directions. The implicit OpenType approach to elements of shaping for vertical and tall group text (as in L2/16-018 section 4) is too unpredictable for usability and this should be addressed (see http://www.unicode.org/L2/L2016/16214-egyptian-ctrl-chars.pdf for possible enhancements to deal with this).

A compromise solution

Actions
Add the EGYPTIAN HIEROGLYPH HORIZONAL JOINER and EGYPTIAN HIEROGLYPH VERTICAL JOINER characters to Unicode 10 on the basis of L2/16-018 but exclude LIGATURE at this first stage. Scope to horizontal text only.

Agree two simple higher level protocols to be used alongside this limited system to enable work to move forward on specialist software support and database encoding in Unicode. One simple HLP for majority everyday use and a superset for specialist requirements (including vertical text, tall groups, and rare quadrats).

Update section 11.4 *Egyptian Hieroglyphs* of the Unicode Standard to use the two control system rather than MdC (as stands at Unicode 9.0). Changes to Unicode data tables and technical notes?

Immediate follow up
Establish specialist Egyptologist requirements clearly and consult with a wider expert user base. Coordinate an extended control scheme with the first expansion to the hieroglyphic repertoire. Start immediately and it should be possible to resolve key issues in months not years. The greatest Unicode limitation for hieratic transcription and many other purposes is a relatively small number of unencoded hieroglyphs rather than control-related issues so both need addressing soon. Prepare consensus update for consideration at UTC meeting starting 31st October.

Benefits
Egyptologists will have the opportunity to understand how hieroglyphic as a Unicode writing system works.

General purpose software (web browsers, word processors, search engines, low level shaping software can be updated to support Egyptian Hieroglyphic as a complex script. Once Egyptian is treated as a complex script in software, additional controls should be technically straightforward to add.

Fonts can implement Egyptian Hieroglyphic as a complex script with basic functionality.

Specialist software can use a simple HLP to enable progression from MdC-level designs and data formats.

Specialist databases (such as Ramses and TLA) can use a simple HLP to enable progression from MdC type formats.

Disadvantages
For the majority of scholars and virtually all casual users, using a system that almost but not quite meets their needs will be a little frustrating until more controls are added.

Web sites such as Wikipedia will probably want to stick with non-text methods for rendering hieroglyphs (e.g. WikiHiero) until a more complete solution is available.