Release Management Group Report to UTC #179

April 17, 2024
Peter Constable & UTC Release Management Group

The Release Management Group ("RMG") is composed of UTC working group leaders and other volunteers working on the process for development of the Unicode 16.0 release and for evolving processes for future Unicode Standard releases.

Unicode 16.0 timeline

A timeline for the development and release of Unicode 16.0 was proposed in L2/23-264 and confirmed by UTC #177 (cf. 177-C1). The following is a summary of key dates:

- 2024-4-2: End of alpha
- 2024-4-25, UTC #179: Finalize beta content
- 2024-5-21: start of beta review
- 2024-7-25, UTC #180: Finalize 16.0 content
- 2024-9-10: Unicode 16.0 release

Unicode 16.0 beta review

The 16.0 beta review period will begin four weeks after UTC #179, on May 21. The RMG together with the Charts WG, CJK, EDC and PAG have begun preparing for this milestone. The beta deliverables will consist of:

- A complete draft of core spec in the Web format (❣️🥳);
- updated code charts;
- all data files;
- a PRI cover page; and
- a Unicode blog post.

RMG recommends the following UTC action:

[179-C] Consensus: The UTC authorizes starting the beta review for Unicode 16.0.

[179-A] Action Item for Peter Constable, RMG: Coordinate the publication of the Unicode 16.0 beta review.

[179-A] Action Item for Ken Whistler, RMG: Prepare a background document for a PRI on the Unicode 16.0 beta review.

[179-A] Action Item for Rick McGowan, UTC: Post the PRI for the Unicode 16.0 beta review, to close July 2, 2024.

Preliminary repertoire for Unicode 17.0

By the end of the Q4 UTC meeting (UTC #181), UTC should have made preliminary decisions on the repertoire for Unicode 17.0. (A draft repertoire for the Unicode 17.0 alpha will be required no later than UTC #182.)

In the meantime, we also know that JTC 1/SC 2 will be meeting in June and will likely be deciding on the next edition for ISO/IEC 10646. While it's too early to know what the next edition might look like (e.g., SC2 could decide on a new edition that doesn't add new repertoire but, rather, establishes a maintenance agency for handing new repertoire), it will certainly be helpful for the 10646 project editor and for Unicode and US delegates to have some idea of what UTC anticipates as repertoire for Unicode 17.0.

With that in mind, RMG has reviewed the pipeline and come up with a preliminary proposal for character and script additions for Unicode 17.0. Key decision criteria for this list included:

- Prioritize characters based on how long they have been in *provisionally assigned* status.
- Prioritize characters in modern usage by language communities.
- Consider proposals for related characters that should progress together.

Based on these factors, our preliminary proposal includes the additions summarized below. Note that some of the characters listed will be proposed for provisional code point assignment at UTC #179; these code points are marked in the table below with an asterisk (*). Also listed is CJK Extension J, which it is anticipated will be proposed for encoding at UTC #180.

RMG is not asking for any UTC action at this time, but invites discussion. We may propose formal approvals for Unicode 17.0 at UTC #180 or UTC #181.

New scripts / blocks

Characters / scripts	Provisional code points	Prior UTC actions
CJK Unified Ideographs Extension J	TBD	
Beria Erfe script	16EBB..16ED3*	
Chisoi script	16D80..16DA9	174-C21
Sidetic script	10940..1095C	174-C17
Tai Yo script	1E6C0..1E6FF	174-C23
Tolong Siki script	11DB0..11DE9	174-C22

Additions to existing scripts/blocks:

Characters / scripts	Provisional code points	Prior UTC actions
Arabic letters, marks, etc.	FBC3..FBD2*, FD90, FD91*, FDC8..FDCE, 10EC5..10EC6, 10EC7*, 10ED0, 10ED1..10ED8*, 10EFA..10EFB	174-C18, 175-C13, 176-C33, 177-C21, 178-C16
Bengali letter Sanskrit ba	09FF	174-C20, 178-C38
Chinese small letters	16FF2..16FF3	178-C27
IPA/phonetic letters & diacritics	1AD0..1ADD, A7CE..A7CF	176-C35, 177-C20, 177-C38, 177-C39
Sharada vowel signs for Kashmiri	11B60..11B67	176-C34
Latin letters	A7D2, A7D4	176-C32
Symbols	2B96, 1CCFA..1CCFC, 1CEBA..1CED0, 1CEE0..1CEF0, 1F777..1F77A, 1F8D0..1F8D8, 1FA54..1FA57, 1FBFA	177-C32, 177-C33, 177-C34, 177-C35, 178-C33, 178-C35
Mongolian letter	1879	178-C30
Tangut components, ideographs	187F8..187FF, 18D09..18D1C, 18D80..18DF3	177-C27, 177-C28

Progression of new UTRs

Historically, UTC would record decisions for progressing a proposed new UTR (UAX, etc.) through successive stages—UTC would approve each of the following:

- Proposed Draft UTR
- Draft UTR
- UTR

These stages are mentioned in the [About Unicode Technical Reports](#) page.

Some years ago, we also introduced posting public review issues for feedback in the Proposed Draft and Draft phases.

Recently, a question was raised whether a *new* PRI was needed when the document advances to Draft status, or if the PRI for the Proposed Draft could be kept open (with revised wording) for the Draft. For example, UTC #178 recorded four actions related to progression of UAX #57, including the following:

> [**178-A29**] **Action Item for** Rick McGowan, UTC: Close PRI #493 Proposed Draft UAX # 57 Unicode Egyptian Hieroglyph Database (Unikemet). [Ref. Section 2b of L2/24-013R]
>
> [**178-A31**] **Action Item for** Rick McGowan: Post a PRI for Draft UAX #57 Unicode Egyptian Hieroglyph Database (Unikemet) for Unicode Version 16.0. [Ref. Section 2b of L2/24-013R]

It has been suggested that the two action items A29 and A31 could be streamlined if PRI #493 were simply kept open.

RMG discussed these questions and propose the following for UTC consideration:
- Keep the existing framework of defined stages—Proposed Draft UTR, Draft UTR, UTR.
- For future cases of new UTRs, allow the PRI for a Proposed Draft to remain open (with appropriate wording changes) as the document advances to Draft stage.
- For proposals that are deemed to be sufficiently mature from the outset, allow these to (formally) start at the Draft UTR stage, by-passing the Proposed Draft stage.

Acknowledgements

In the past, one volunteer (Lisa Moore) would compile the list of people that should be acknowledged as contributors for each new version of the standard. Sometimes, this could have involved reconstructing, at some point late in the cycle, who contributed what.

Since the UTC process has evolved to have more preparatory work done outside of UTC proper and in the various working groups, it has been harder for one volunteer or even one working group (e.g., EDC) to track contributors for the release.

To address this issue, RMG has created a standing issue in its issue-tracking repo, which the various working group leaders have access to. RMG has asked the leadership for each working group to capture in that one place contributions that merit acknowledgement. This distributes the workload to the people that have the best awareness of contributions. It also makes it easy for leadership to regularly capture contributions throughout the development cycle, rather than having to do a historical review later in the cycle.

Related to this, acknowledgements in UTRs (UAXs, UTSs) have been handled inconsistently. This was discussed by EDC, who have adopted a policy for consistent handling of acknowledgements in UTRs.

Process development

Each Unicode release involves a long list of artifacts (core spec content, UAX pages, data files, website cover material, etc.) that need to be collected and deployed. The various artifacts have been handled using multiple, different workflows and several manual steps. Some of those steps have historically been done outside of any Unicode infrastructure or with custom tools (e.g., on Ken Whistler's personal machine with his own tooling); at present, we would not be able to reconstruct such workflows if the individuals that have handled those were no longer available.

RMG has begun to look at how deployment of artifacts might be made more robust and more streamlined in the future. We do not anticipate putting any new workflow into operation for Unicode 16.0, but would aim to make changes for Unicode 17.0.