



Proposed Draft Unicode Technical Report #54

UNICODE® MONGOLIAN 12.1 BASELINE

Authors	Ken Whistler (ken@unicode.org)
Date	2019-06-06
This Version	http://www.unicode.org/reports/tr54/tr54-1.html
Previous Version	
Latest Version	http://www.unicode.org/reports/tr54/
Latest Proposed Update	http://www.unicode.org/reports/tr54/proposed.html
Revision	1

Summary

This technical report documents the state of the Unicode Mongolian code chart, names list, and glyph variants as of Unicode 12.1.

Status

*This is a **draft** document which may be updated, replaced, or superseded by other documents at any time. Publication does not imply endorsement by the Unicode Consortium. This is not a stable document; it is inappropriate to cite this document as other than a work in progress.*

A Unicode Technical Report (UTR) contains informative material. Conformance to the Unicode Standard does not imply conformance to any UTR. Other specifications, however, are free to make normative references to a UTR.

Please submit corrigenda and other comments with the online reporting form [[Feedback](#)]. Related information that is useful in understanding this document is found in the [References](#). For the latest version of the Unicode Standard see [[Unicode](#)]. For a list of current Unicode Technical Reports see [[Reports](#)]. For more information about versions of the Unicode Standard, see [[Versions](#)].

Contents

- 1 Overview
 - 2 Background
 - 2 Mongolian Version 12.1 Code Chart
 - References
 - Modifications
-

1 Overview

After the publication of Unicode 12.1, the Unicode Technical Committee decided that it would be best to split the problem of documentation of the details of glyph variation in Mongolian from the documentation of the code points, names, and basic reference glyphs for Mongolian characters in the Unicode code charts.

This technical report preserves the state of code charts for Mongolian as of Unicode 12.1, the last version which printed all the variant information interspersed with the rest of the basic code chart information. This report should serve as an easily accessible baseline for the development of alternative documentation vehicles for Mongolian glyph variation, which will be more complete, better aligned with a fully explained text model for Mongolian, and easier to maintain directly than Unicode code charts.

2 Background

There are a number of strong reasons to separate the documentation of glyph variation for Mongolian from the basic code charts.

There are two major classes of glyph variation impacting the rendering and display of Mongolian text:

- Positional Variants
- Standardized Variation Sequences

The positional variants are changes in the shapes of Mongolian letters, depending on their relative position with a cursively connected sequence of letters. In particular, letter forms may change shape in initial, medial, and/or final positions. The standardized variation sequences involve specific sequences of a letter plus a one of several dedicated Mongolian Free Variation Selector (FVS) characters. Each defined combination of letter plus FVS has a particular expected shape outcome. For both positional variants and standardized variation sequences, in many instances, the shape may be influenced by further details of the morphological context for the word.

The Mongolian code charts have gradually gotten more complex over the years, in an attempt to convey all this shaping information. As of Version 12.1, all of the positional variants and all of the standardized variation sequences were displayed together in the Mongolian code chart. The net effect has turned out not to be particularly clear or helpful—instead the result is to overwhelm the user of the code chart with a mass of glyph variant detail, without sufficient information or structure to make sense of all this detail. Separation of this information into structured alternatives to the basic code charts can result in much more comprehensible outcomes.

A second major concern is that the Mongolian script encoding encompasses more than just the writing system for modern Mongolian (Hudum). It also includes repertoire to cover Todo, Sibe, Manchu, and Ali Gali (Sanskrit). The code chart necessarily includes all of the

repertoire for all of these writing systems together in one chart. The inclusion of all repertoire in one chart makes it very difficult to distinguish the relevant glyph variation which applies to each writing system.

Third, the tooling which is used to format the PDF file for the Unicode code charts is very complex software that was not optimized specifically for any one script, and certainly not for the details of Mongolian glyph variants. It can be very frustrating for people who wish to improve the documentation of the Mongolian text model or to improve that model, when they run into the tooling barriers which make it difficult to easily adjust the display of glyph variants. Separation of the glyph variant information into documents that can be maintained with normal editorial tools, instead of complicated, proprietary, code chart formatting software, would make it possible to be much more responsive to reports of problems in the documentation and in the Mongolian text model.

Finally, the Unicode code charts are only updated once with each new release of the Unicode Standard, and each update is therefore closely tied to the Unicode release cycle. Separation of the glyph variant information and documentation of all the associated contextual rules and their interaction with the Mongolian text model, from the production of versioned code charts would also make it possible to update this information much more quickly. Documentation types such as Unicode Technical Reports and Unicode Technical Notes are not tied to the Unicode release cycle, and thus would be more appropriate vehicles for quick turnaround of documentation and model improvements.

3 Mongolian Version 12.1 Code Chart

The Mongolian code chart for Version 12.1 of the Unicode Standard is available at the following link:

[U1800_v12_1_20190426.pdf](#)

That chart is identical to what was published for the Mongolian block as of Version 12.1, including the cover sheet, as well as the list of all variant forms appended after the formatted names list. That chart is also accessible in the archived complete code charts for the entire Unicode 12.1 standard, but for Mongolian reference purposes, it is much easier to refer here to this single code chart for the Mongolian block.

The code chart for the Mongolian Supplement block (U+11660..U+1167F) contains no information about glyph variants for normal Mongolian letters, so is not included here. The maintenance of that code chart is not a particular issue of concern for the details of the Mongolian text model.

References

TBD

Modifications

The following summarizes modifications from the previous revision of this document.

Revision 1:

- Initial proposed draft Unicode Technical Report.

© 2019 Unicode, Inc. All Rights Reserved. The Unicode Consortium makes no expressed or implied warranty of any kind, and assumes no liability for errors or omissions. No liability is assumed for incidental and consequential damages in connection with or arising out of the use of the information or programs contained or accompanying this technical report. The Unicode **Terms of Use** apply.

Unicode and the Unicode logo are trademarks of Unicode, Inc., and are registered in some jurisdictions.