



Proposed Update Unicode Standard Annex #41

COMMON REFERENCES FOR UNICODE STANDARD ANNEXES

Version	Unicode 6.0.0 draft2
Author	Magda Danish
Date	2010-04-09
This Version	http://www.unicode.org/reports/tr41/tr41-6.html
Previous Version	http://www.unicode.org/reports/tr41/tr41-5.html
Latest Version	http://www.unicode.org/reports/tr41/
Latest Proposed Update	http://www.unicode.org/reports/tr41/proposed.html
Revision	6

Summary

This annex presents a common set of references for the Unicode Standard Annexes.

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*This is a **draft** document which may be updated, replaced, or superseded by other documents at any time.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.*

This document has been reviewed by Unicode members and other interested parties, and has been approved for publication by the Unicode Consortium. This is a stable document and may be used as reference material or cited as a normative reference by other specifications.

A Unicode Standard Annex (UAX) forms an integral part of the Unicode Standard, but is published online as a separate document. The Unicode Standard may require conformance to normative content in a Unicode Standard Annex, if so specified in the Conformance chapter of that version of the Unicode Standard. The version number of a UAX document corresponds to the version of the Unicode Standard of which it forms a part.

Please submit corrigenda and other comments with the online reporting form

[[Feedback](#)]. Related information that is useful in understanding this annex is found in Unicode Standard Annex #41, “[Common References for Unicode Standard Annexes](#).” 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](#)]. For any errata which may apply to this annex, see [[Errata](#)].

Contents

- 1 [References to Publications by the Unicode Consortium](#)
 - 2 [References to Other Standards](#)
 - 3 [Other References](#)
-

1 References to Publications by the Unicode Consortium

Publications may be listed more than once under different headings.

- [Blocks] Blocks data file
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/Blocks.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/Blocks.txt>
- [Charts] Online Code Charts
<http://www.unicode.org/charts/>
An index to character names with links to the corresponding chart is found at
<http://www.unicode.org/charts/charindex.html>
- [Charts14] Charts for the test files
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/auxiliary/LineBreakTest.html>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/LineBreakTest.html>
- [Charts15] Normalization Charts
<http://www.unicode.org/reports/tr15/charts>
- [Charts29] Charts for the test files
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/auxiliary/GraphemeBreakTest.html>
<http://www.unicode.org/Public/UNIDATA/auxiliary/WordBreakTest.html>
<http://www.unicode.org/Public/UNIDATA/auxiliary/SentenceBreakTest.html>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/GraphemeBreakTest.html>
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/WordBreakTest.html>
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/SentenceBreakTest.html>

- [/SentenceBreakTest.html](#)
- [CLDR] Unicode Locales Project (Unicode Common Locale Data Repository)
<http://www.unicode.org/cldr/>
- [Code9] Reference code implementing Unicode Bidirectional Algorithm
For Java reference code, see:
<http://www.unicode.org/Public/PROGRAMS/BidiReferenceJava>
For C/C++ reference code, see:
<http://www.unicode.org/Public/PROGRAMS/BidiReferenceCpp>
- [Code14] Sample code implementing Unicode Line Breaking Algorithm using a pair table
<http://www.unicode.org/Public/PROGRAMS/LineBreakSampleCpp/>
Contains the code samples shown in UAX #14 together with driver code.
- [Corrections] Normalization Corrections
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/NormalizationCorrections.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/NormalizationCorrections.txt>
- [Corrigendum1] Corrigendum #1: *UTF-8 Shortest Form*
<http://www.unicode.org/versions/corrigendum1.html>
- [Corrigendum2] Corrigendum #2: *Yod with Hiriq Normalization*
<http://www.unicode.org/versions/corrigendum2.html>
- [Corrigendum3] Corrigendum #3: *U+F951 Normalization*
<http://www.unicode.org/versions/corrigendum3.html>
- [Corrigendum4] Corrigendum #4: *Five CJK Canonical Mapping Errors*
<http://www.unicode.org/versions/corrigendum4.html>
- [Corrigendum5] Corrigendum #5: *Normalization Idempotency*
<http://www.unicode.org/versions/corrigendum5.html>
- [Corrigendum6] Corrigendum #6: *Bidi Mirroring*
<http://www.unicode.org/versions/corrigendum6.html>
- [Corrigendum7] Corrigendum #7: *UAX #14, Unicode Line Breaking Algorithm, rule LB8*
<http://www.unicode.org/versions/corrigendum7.html>
- [Data9] Bidi Mirroring
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/BidiMirroring.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/BidiMirroring.txt>
- [Data11] East Asian Width property data file
For the latest version, see:

- <http://www.unicode.org/Public/UNIDATA/EastAsianWidth.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/EastAsianWidth.txt>
- [Data14] Line Break property data file
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/LineBreak.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/LineBreak.txt>
- [Data24] Scripts data file
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/Scripts.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/Scripts.txt>
- [Data34] Named Sequences data file
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/NamedSequences.txt>
For the 5.2.0 version see:
<http://www.unicode.org/Public/5.2.0/ucd/NamedSequences.txt>
- [DataProv] Provisional Named Sequences data file
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/NamedSequencesProv.txt>
For the 5.2.0 version see:
<http://www.unicode.org/Public/5.2.0/ucd/NamedSequencesProv.txt>
- [DerivedBIDI] Derived Bidi Properties
For the latest version see:
<http://www.unicode.org/Public/UNIDATA/extracted/DerivedBidiClass.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/extracted/DerivedBidiClass.txt>
- [Errata] Updates and Errata
<http://www.unicode.org/errata>
- [Exclusions] Composition Exclusion Table
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/CompositionExclusions.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/CompositionExclusions.txt>
- [FAQ] Unicode Frequently Asked Questions
<http://www.unicode.org/faq/>
For answers to common questions on technical issues.

- [Feedback] Reporting Form
<http://www.unicode.org/reporting.html>
For reporting errors and requesting information online.
- [Glossary] Unicode Glossary
<http://www.unicode.org/glossary/>
For explanations of terminology used in this and other documents.
- [HangulST] Hangul Syllable Types
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/HangulSyllableType.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/HangulSyllableType.txt>
- [NormProps] Derived Normalization Properties
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/DerivedNormalizationProps.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/DerivedNormalizationProps.txt>
- [Policies] Unicode Policies
<http://www.unicode.org/policies/policies.html>
- [Props] Unicode Text Segmentation Property Data:
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/auxiliary/GraphemeBreakProperty.txt>
<http://www.unicode.org/Public/UNIDATA/auxiliary/WordBreakProperty.txt>
<http://www.unicode.org/Public/UNIDATA/auxiliary/SentenceBreakProperty.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/GraphemeBreakProperty.txt>
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/WordBreakProperty.txt>
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/SentenceBreakProperty.txt>
- [PropValue] Property Value Aliases data file
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/PropertyValueAliases.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/PropertyValueAliases.txt>
- [Reports] Unicode Technical Reports
<http://www.unicode.org/reports/>
For information on the status and development process for technical reports, and for a list of technical reports.

- [Sample] Sample Normalizer code
<http://www.unicode.org/reports/tr15/Normalizer.html>
- [Stability] Unicode Consortium Stability Policies
http://www.unicode.org/policies/stability_policy.html
- [Tests14] Test data:
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/auxiliary/LineBreakTest.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/LineBreakTest.txt>
- [Tests15] Normalization Conformance Test
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/NormalizationTest.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/NormalizationTest.txt>
- [Tests29] Test data:
For the latest version, see:
<http://www.unicode.org/Public/UNIDATA/auxiliary/GraphemeBreakTest.txt>
<http://www.unicode.org/Public/UNIDATA/auxiliary/WordBreakTest.txt>
<http://www.unicode.org/Public/UNIDATA/auxiliary/SentenceBreakTest.txt>
For the 5.2.0 version, see:
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/GraphemeBreakTest.txt>
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/WordBreakTest.txt>
<http://www.unicode.org/Public/5.2.0/ucd/auxiliary/SentenceBreakTest.txt>
- [UAX9] UAX #9: *Unicode Bidirectional Algorithm*
<http://www.unicode.org/reports/tr9/>
- [UAX11] UAX #11: *East Asian Width*
<http://www.unicode.org/reports/tr11/>
- [UAX14] UAX #14: *Unicode Line Breaking Algorithm*
<http://www.unicode.org/reports/tr14/>
- [UAX15] UAX #15: *Unicode Normalization Forms*
<http://www.unicode.org/reports/tr15/>
- [UAX24] UAX #24: *Unicode Script Property*
<http://www.unicode.org/reports/tr24/>
- [UAX29] UAX #29: *Unicode Text Segmentation*
<http://www.unicode.org/reports/tr29/>
- [UAX31] UAX #31: *Unicode Identifier and Pattern Syntax*
<http://www.unicode.org/reports/tr31/>
- [UAX34] UAX #34: *Unicode Named Character Sequences*
<http://www.unicode.org/reports/tr34/>

- [UAX38] UAX #38: *Unicode Han Database (Unihan)*
<http://www.unicode.org/reports/tr38/>
- [UAX41] UAX #41: *Common References for Unicode Standard Annexes*
<http://www.unicode.org/reports/tr41/>
- [UAX42] UAX #42: *Unicode Character Database in XML*
<http://www.unicode.org/reports/tr42/>
- [UAX44] UAX #44: *Unicode Character Database*
<http://www.unicode.org/reports/tr44/>
- [UCA] UTS #10: *Unicode Collation Algorithm*
<http://www.unicode.org/reports/tr10/>
- [UCD] Unicode Character Database
<http://www.unicode.org/ucd/>
For detailed documentation about the Unicode Character Database, see Unicode Standard Annex #44: Unicode Character Database
<http://www.unicode.org/reports/tr44/>
- [Unicode] The Unicode Standard
For the latest version, see:
<http://www.unicode.org/versions/latest/>
For the 5.2.0 version, see:
<http://www.unicode.org/versions/Unicode5.2.0/>
- [Unicode3.0] The Unicode Consortium. *The Unicode Standard, Version 3.0* (Reading, MA, Addison–Wesley, 2000. ISBN 0–201–61633–5).
- [Unicode3.1] The Unicode Consortium. The Unicode Standard, Version 3.1.0, defined by: *The Unicode Standard, Version 3.0* (Reading, MA, Addison–Wesley, 2000. ISBN 0–201–61633–5), as amended by the *Unicode Standard Annex #27: Unicode 3.1*
<http://www.unicode.org/reports/tr27/>
- [Unicode3.2] The Unicode Consortium. The Unicode Standard, Version 3.2.0, defined by: *The Unicode Standard, Version 3.0* (Reading, MA, Addison–Wesley, 2000. ISBN 0–201–61633–5), as amended by the *Unicode Standard Annex #27: Unicode 3.1* and the *Unicode Standard Annex #28: Unicode 3.2*
<http://www.unicode.org/reports/tr28/>
- [Unicode4.0] The Unicode Consortium. *The Unicode Standard, Version 4.0* (Boston, MA, Addison–Wesley, 2003. ISBN 0–321–18578–1).
- [Unicode4.0.1] The Unicode Consortium. The Unicode Standard, Version 4.0.1, defined by: *The Unicode Standard, Version 4.0* (Boston, MA, Addison–Wesley, 2003. ISBN 0–321–18578–1), as amended by *Unicode 4.0.1*
<http://www.unicode.org/versions/Unicode4.0.1/>
- [Unicode4.1] The Unicode Consortium. The Unicode Standard, Version 4.1.0, defined by: *The Unicode Standard, Version 4.0* (Boston, MA, Addison–Wesley, 2003. ISBN

- 0-321-18578-1), as amended by [Unicode 4.0.1](#) and by [Unicode 4.1.0](#)
<http://www.unicode.org/versions/Unicode4.1.0/>
- [Unicode5.0] The Unicode Consortium. [The Unicode Standard, Version 5.0](#) (Boston, MA, Addison-Wesley, 2007. ISBN 0-321-48091-0).
- [Unicode5.1] The Unicode Consortium. The Unicode Standard, Version 5.1.0, defined by: The Unicode Standard, Version 5.0 (Boston, MA, Addison-Wesley, 2007. ISBN 0-321-48091-0), as amended by [Unicode 5.1.0](#)
- [Unicode5.2] The Unicode Consortium. The Unicode Standard, Version 5.2.0, defined by: [The Unicode Standard, Version 5.2](#) (Mountain View, CA: The Unicode Consortium, 2009. ISBN 978-1-936213-00-9)
- [UTC] Unicode Technical Committee
<http://www.unicode.org/consortium/utc.html>
- [UTN5] UTN #5: *Canonical Equivalences in Applications*
<http://www.unicode.org/notes/tn5>
- [UTR17] UTR #17: *Unicode Character Encoding Model*
<http://www.unicode.org/reports/tr17/>
- [UTR20] UTR # 20: *Unicode in XML and other Markup Languages*
<http://www.unicode.org/reports/tr20/>
- [UTR23] UTR # 23: *The Unicode Character Property Model*
<http://www.unicode.org/reports/tr23/>
- [UTR25] UTR # 25: *Unicode Support for Mathematics*
<http://www.unicode.org/reports/tr25/>
- [UTR33] UTR # 33: *Unicode Conformance Model*
<http://www.unicode.org/reports/tr33/>
- [UTR36] UTR #36: *Unicode Security Considerations*
<http://www.unicode.org/reports/tr36/>
- [UTR45] UTR #45: *U-Source Ideographs*
<http://www.unicode.org/reports/tr45/>
- [UTS6] UTS #6: *A Standard Compression Scheme for Unicode*
<http://www.unicode.org/reports/tr6/>
- [UTS10] UTS #10: *Unicode Collation Algorithm (UCA)*
<http://www.unicode.org/reports/tr10/>
- [UTS18] UTS #18: *Unicode Regular Expressions*
<http://www.unicode.org/reports/tr18/>
- [UTS22] UTS #22: *Unicode Character Mapping Markup Language*
<http://www.unicode.org/reports/tr22/>
- [UTS35] UTS #35: *Unicode Locale Data Markup Language (LDML)*
<http://www.unicode.org/reports/tr35/>
- [UTS37] UTS #37: *Unicode Ideographic Variation Database*
<http://www.unicode.org/reports/tr37/>
- [UTS39] UTS #39: *Unicode Security Mechanisms*
<http://www.unicode.org/reports/tr39/>
- [Versions] Versions of the Unicode Standard
<http://www.unicode.org/versions/>
For information on version numbering, and citing and

referencing the Unicode Standard, the Unicode Character Database, and Unicode Technical Reports.

2 References to Other Standards

- [10646] International Organization for Standardization. *Information Technology—Universal Multiple-Octet Coded Character Set (UCS)*. (ISO/IEC 10646:2003). *For availability, see:*
<http://www.iso.org>
- [ISO15924] ISO 15924: *Code for the Representation of Names of Scripts*
<http://www.unicode.org/iso15924/>
- [ISO19757] ISO/IEC 19757-2:2003 – Information technology – Document Schema Definition Language (DSDL) – Part 2: Regular-grammar-based validation – RELAX NG
Available at <http://standards.iso.org/ittf/PubliclyAvailableStandards/>.
- [ISO19757 Amd 1] ISO/IEC 19757-2:2003 – Information technology – Document Schema Definition Language (DSDL) – Part 2: Regular-grammar-based validation – RELAX NG – Amendment 1: Compact Syntax
Available at <http://standards.iso.org/ittf/PubliclyAvailableStandards/>.
- [JIS] JIS X 4051-1995. *Line Composition Rules for Japanese Documents*.
(『日本語文書の行組版方法』) Japanese Standards Association. 1995.
- [XML] Extensible Markup Language (XML) 1.0, 5th Edition or later
<http://www.w3.org/TR/2008/REC-xml-20081126/>

3 Other References

- [Cedar97] Cy Cedar, David Veintimilla, Michel Suignard, and Asmus Freytag, *Report from the Trenches: Microsoft Publisher goes Unicode*. Proceedings of the Eleventh International Unicode Conference, San Jose, CA, 1997.
- [CharLint] Charlint—A Character Normalization Tool
<http://www.w3.org/International/charlint/>
- [CharMod] Martin J. Dürst, François Yergeau, Richard Ishida, Misha Wolf, and Tex Texin, [W3C Character Model for the World Wide Web](http://www.w3.org/TR/charmod/), (See <http://www.w3.org/TR/charmod/>)
- [CharNorm] Martin J. Dürst, François Yergeau, Richard Ishida, Misha Wolf, Tex Texin, and Addison Phillips, [Character Model for the World Wide Web 1.0: Normalization](http://www.w3.org/TR/charmod-norm/), W3C Working Draft. (See [http://www.w3.org/TR/charmod-norm.](http://www.w3.org/TR/charmod-norm/))
- [CharReq] Martin J. Dürst, [Requirements for String Identity Matching and String Indexing](http://www.w3.org/TR/WD-charreq/), W3C Working Draft. (See [http://www.w3.org/TR/WD-charreq.](http://www.w3.org/TR/WD-charreq/))

- [Knuth78] Donald E. Knuth and Michael F. Plass, *Breaking Lines into Paragraphs*, republished in *Digital Typography*, CSLI 78 (Stanford, California: CLSI Publications 1997).
- [Suign98] Michel Suignard, *Worldwide Typography and How to Apply JIS X 4051-1995 to Unicode*. Proceedings of the Twelfth International Unicode/ISO 10646 Conference, Tokyo, Japan, 1998.
- [TEX] Donald E. Knuth, *TEX, the Program*, Volume B of *Computers & Typesetting* (Reading, MA, Addison-Wesley, 1986).
-

Copyright © 2001-2010 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.