



Proposed Update Unicode Standard Annex #41

COMMON REFERENCES FOR UNICODE STANDARD ANNEXES

Version	Unicode 8.0.0 (draft 3)
Editors	Laurențiu Iancu, Rick McGowan
Date	2014-12-03
This Version	http://www.unicode.org/reports/tr41/tr41-16.html
Previous Version	http://www.unicode.org/reports/tr41/tr41-15.html
Latest Version	http://www.unicode.org/reports/tr41/
Latest Proposed Update	http://www.unicode.org/reports/tr41/proposed.html
Revision	16

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*This is a **draft** document which may be updated, replaced, or superseded by other documents at any time.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

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](#)]. 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](#)] Character Block Property Data File
Latest version:
<http://www.unicode.org/Public/UCD/latest/ucd/Blocks.txt>
Version 8.0.0:
<http://www.unicode.org/Public/8.0.0/ucd/Blocks.txt>
- [[Charts](#)] Character Code Charts
Latest version:
<http://www.unicode.org/charts/>
Index of character names with links to the corresponding code charts:
<http://www.unicode.org/charts/charindex.html>
- [[Charts](#)14] Charts for the Unicode Line Breaking Algorithm Test Files
Latest version:
<http://www.unicode.org/Public/UCD/latest/ucd/auxiliary/LineBreakTest.html>
Version 8.0.0:
<http://www.unicode.org/Public/8.0.0/ucd/auxiliary/LineBreakTest.html>
- [[Charts](#)15] Normalization Charts
Latest version:
<http://www.unicode.org/charts/normalization/>

[Charts29]

Charts for the Unicode Text Segmentation Test Files

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/auxiliary/GraphemeBreakTest.html><http://www.unicode.org/Public/UCD/latest/ucd/auxiliary/WordBreakTest.html><http://www.unicode.org/Public/UCD/latest/ucd/auxiliary/SentenceBreakTest.html>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/auxiliary/GraphemeBreakTest.html><http://www.unicode.org/Public/8.0.0/ucd/auxiliary/WordBreakTest.html><http://www.unicode.org/Public/8.0.0/ucd/auxiliary/SentenceBreakTest.html>**[CLDR]**

Unicode Locales Project (Unicode Common Locale Data Repository)

<http://cldr.unicode.org/>**[Code9]**

Reference Implementations of the Unicode Bidirectional Algorithm

C reference code:

<http://www.unicode.org/Public/PROGRAMS/BidiReferenceC/>

Java reference code:

<http://www.unicode.org/Public/PROGRAMS/BidiReferenceJava/>**[Code14]**

Sample Implementation of the Unicode Line Breaking Algorithm

<http://www.unicode.org/Public/PROGRAMS/LineBreakSampleCpp/>**[Corrections]**

Normalization Corrections Data File

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/NormalizationCorrections.txt>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0>

[/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>

[[Corrigendum8](#)] **Corrigendum #8:** Bidi_Class Fix for U+070F Syriac Abbreviation Mark

<http://www.unicode.org/versions/corrigendum8.html>

[[Corrigendum9](#)] **Corrigendum #9:** Clarification About Noncharacters

<http://www.unicode.org/versions/corrigendum9.html>

[[Data9](#)] **Unicode Bidirectional Algorithm Property Data Files**

Latest version:

<http://www.unicode.org/Public/UCD/latest>

[/ucd/BidiMirroring.txt](#)

<http://www.unicode.org/Public/UCD/latest>

[/ucd/BidiBrackets.txt](#)

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0>

[/ucd/BidiMirroring.txt](#)

<http://www.unicode.org/Public/8.0.0>

[/ucd/BidiBrackets.txt](#)

[[Data11](#)] **East Asian Width Property Data File**

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/EastAsianWidth.txt>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/EastAsianWidth.txt>

[[Data14](#)] Unicode Line Breaking Algorithm Property Data File

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/LineBreak.txt>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/LineBreak.txt>

[[Data24](#)] Unicode Script Property Data File

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/Scripts.txt>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/Scripts.txt>

[[Data34](#)] Unicode Named Character Sequences Data File

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/NamedSequences.txt>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/NamedSequences.txt>

[[Data45](#)] U-Source Ideographs Data File

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/USourceData.txt>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/USourceData.txt>

[[DataProv](#)] Provisional Named Sequences Data File

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/NamedSequencesProv.txt>

- Version 8.0.0:**
<http://www.unicode.org/Public/8.0.0/ucd/NamedSequencesProv.txt>
- [Demo9]** Online Demo of a reference implementation of the Unicode Bidirectional Algorithm
<http://www.unicode.org/cldr/utility/bidi.jsp>
- [DerivedBIDI]** Derived Bidirectional Type Property Data File
Latest version:
<http://www.unicode.org/Public/UCD/latest/ucd/extracted/DerivedBidiClass.txt>
Version 8.0.0:
<http://www.unicode.org/Public/8.0.0/ucd/extracted/DerivedBidiClass.txt>
- [Errata]** Updates and Errata
<http://www.unicode.org/errata>
- [Exclusions]** Composition Exclusion Table
Latest version:
<http://www.unicode.org/Public/UCD/latest/ucd/CompositionExclusions.txt>
Version 8.0.0:
<http://www.unicode.org/Public/8.0.0/ucd/CompositionExclusions.txt>
- [FAQ]** Unicode Frequently Asked Questions
Answers to common questions on technical issues:
<http://www.unicode.org/faq/>
- [Feedback]** Contact Form
Error reporting and information requests:
<http://www.unicode.org/reporting.html>
- [Glossary]** Glossary of Unicode Terms
<http://www.unicode.org/glossary/>
For explanations of terminology used in this and other documents.
- [Glyphs45]** U-Source Ideographs Glyph Table
Latest version:
<http://www.unicode.org/Public/UCD/latest>

[/ucd/USourceGlyphs.pdf](#)

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0>

[/ucd/USourceGlyphs.pdf](#)

[[HangulST](#)]

Hangul Syllable Type Property Data File

Latest version:

<http://www.unicode.org/Public/UCD/latest>

[/ucd/HangulSyllableType.txt](#)

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0>

[/ucd/HangulSyllableType.txt](#)

[[NormProps](#)]

Derived Normalization Properties Data File

Latest version:

<http://www.unicode.org/Public/UCD/latest>

[/ucd/DerivedNormalizationProps.txt](#)

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0>

[/ucd/DerivedNormalizationProps.txt](#)

[[Policies](#)]

Unicode Consortium Policies

<http://www.unicode.org/policies/policies.html>

[[Props](#)]

Unicode Text Segmentation Property Data Files

Latest version:

<http://www.unicode.org/Public/UCD/latest>

[/ucd/auxiliary/GraphemeBreakProperty.txt](#)

<http://www.unicode.org/Public/UCD/latest>

[/ucd/auxiliary/WordBreakProperty.txt](#)

<http://www.unicode.org/Public/UCD/latest>

[/ucd/auxiliary/SentenceBreakProperty.txt](#)

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/auxiliary>

[/GraphemeBreakProperty.txt](#)

<http://www.unicode.org/Public/8.0.0/ucd/auxiliary>

[/WordBreakProperty.txt](#)

<http://www.unicode.org/Public/8.0.0/ucd/auxiliary>

[/SentenceBreakProperty.txt](#)

- [[PropValue](#)] Property Value Aliases Data File
Latest version:
<http://www.unicode.org/Public/UCD/latest/ucd/PropertyValueAliases.txt>
Version 8.0.0:
<http://www.unicode.org/Public/8.0.0/ucd/PropertyValueAliases.txt>
- [[Reports](#)] Unicode Technical Reports
List of Unicode Standard Annexes, Technical Standards, and 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.
- [[Stability](#)] Unicode Character Encoding Stability Policy
http://www.unicode.org/policies/stability_policy.html
- [[Tests9](#)] Unicode Bidirectional Algorithm Test Data [File](#)
Latest version:
<http://www.unicode.org/Public/UCD/latest/ucd/BidiTest.txt>
<http://www.unicode.org/Public/UCD/latest/ucd/BidiCharacterTest.txt>
Version 8.0.0:
<http://www.unicode.org/Public/8.0.0/ucd/BidiTest.txt>
<http://www.unicode.org/Public/8.0.0/ucd/BidiCharacterTest.txt>
- [[Tests14](#)] Unicode Line Breaking Algorithm Test Data [File](#)
Latest version:
<http://www.unicode.org/Public/UCD/latest/ucd/auxiliary/LineBreakTest.txt>
Version 8.0.0:
<http://www.unicode.org/Public/8.0.0/ucd/auxiliary/LineBreakTest.txt>
- [[Tests15](#)] Unicode Normalization Forms Test Data [File](#)
Latest version:
<http://www.unicode.org/Public/UCD/latest/ucd/NormalizationTest.txt>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/NormalizationTest.txt>

[[Tests29](#)]

Unicode Text Segmentation Test Data [Files](#)

Latest version:

<http://www.unicode.org/Public/UCD/latest/ucd/auxiliary/GraphemeBreakTest.txt>

<http://www.unicode.org/Public/UCD/latest/ucd/auxiliary/WordBreakTest.txt>

<http://www.unicode.org/Public/UCD/latest/ucd/auxiliary/SentenceBreakTest.txt>

Version 8.0.0:

<http://www.unicode.org/Public/8.0.0/ucd/auxiliary/GraphemeBreakTest.txt>

<http://www.unicode.org/Public/8.0.0/ucd/auxiliary/WordBreakTest.txt>

<http://www.unicode.org/Public/8.0.0/ucd/auxiliary/SentenceBreakTest.txt>

Review Note: Prior to final release, links to the 8.0.0 versions of UAXes resolve to the Proposed Updates of the UAXes, if they exist.

[[UAX9](#)]

Unicode Standard Annex #9: Unicode Bidirectional Algorithm

Latest version:

<http://www.unicode.org/reports/tr9/>

Version 8.0.0:

<http://www.unicode.org/reports/tr9/tr9-32.html>

[[UAX11](#)]

Unicode Standard Annex #11: East Asian Width

Latest version:

<http://www.unicode.org/reports/tr11/>

Version 8.0.0:

<http://www.unicode.org/reports/tr11/tr11-29.html>

[[UAX14](#)]

Unicode Standard Annex #14: Unicode Line Breaking Algorithm

Latest version:

<http://www.unicode.org/reports/tr14/>

Version 8.0.0:

- <http://www.unicode.org/reports/tr14/tr14-34.html>
- [UAX15] Unicode Standard Annex #15: Unicode Normalization Forms
Latest version:
<http://www.unicode.org/reports/tr15/>
Version 8.0.0:
<http://www.unicode.org/reports/tr15/tr15-42.html>
- [UAX24] Unicode Standard Annex #24: Unicode Script Property
Latest version:
<http://www.unicode.org/reports/tr24/>
Version 8.0.0:
<http://www.unicode.org/reports/tr24/tr24-23.html>
- [UAX29] Unicode Standard Annex #29: Unicode Text Segmentation
Latest version:
<http://www.unicode.org/reports/tr29/>
Version 8.0.0:
<http://www.unicode.org/reports/tr29/tr29-26.html>
- [UAX31] Unicode Standard Annex #31: Unicode Identifier and Pattern Syntax
Latest version:
<http://www.unicode.org/reports/tr31/>
Version 8.0.0:
<http://www.unicode.org/reports/tr31/tr31-22.html>
- [UAX34] Unicode Standard Annex #34: Unicode Named Character Sequences
Latest version:
<http://www.unicode.org/reports/tr34/>
Version 8.0.0:
<http://www.unicode.org/reports/tr34/tr34-20.html>
- [UAX38] Unicode Standard Annex #38: Unicode Han Database (UniHan)
Latest version:
<http://www.unicode.org/reports/tr38/>
Version 8.0.0:

- <http://www.unicode.org/reports/tr38/tr38-18.html>
- [[UAX41](#)] Unicode Standard Annex #41: Common References for Unicode Standard Annexes
Latest version:
<http://www.unicode.org/reports/tr41/>
Version 8.0.0:
<http://www.unicode.org/reports/tr41/tr41-16.html>
- [[UAX42](#)] Unicode Standard Annex #42: Unicode Character Database in XML
Latest version:
<http://www.unicode.org/reports/tr42/>
Version 8.0.0:
<http://www.unicode.org/reports/tr42/tr42-16.html>
- [[UAX44](#)] Unicode Standard Annex #44: Unicode Character Database
Latest version:
<http://www.unicode.org/reports/tr44/>
Version 8.0.0:
<http://www.unicode.org/reports/tr44/tr44-15.html>
- [[UAX45](#)] Unicode Standard Annex #45: U-Source Ideographs
Latest version:
<http://www.unicode.org/reports/tr45/>
Version 8.0.0:
<http://www.unicode.org/reports/tr45/tr45-13.html>
- [[UCD](#)] About the Unicode Character Database
<http://www.unicode.org/ucd/>
For detailed documentation, see [[UAX44](#)].
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
Latest version:
<http://www.unicode.org/versions/latest/>
Version 8.0.0:

<http://www.unicode.org/versions/Unicode8.0.0/>

[[Unicode3.0](#)] The Unicode Consortium, The Unicode Standard, Version 3.0.0, defined by: The Unicode Standard, Version 3.0 (Reading, MA: Addison–Wesley, 2000. ISBN 0–201–61633–5), <http://www.unicode.org/versions/Unicode3.0.0/>

[[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 4.0.0, defined by: The Unicode Standard, Version 4.0 (Boston, MA: Addison–Wesley, 2003. ISBN 0–321–18578–1), <http://www.unicode.org/versions/Unicode4.0.0/>

[[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 5.0.0, defined by: The Unicode Standard, Version 5.0 (Boston, MA: Addison-Wesley, 2007. ISBN 0-321-48091-0), <http://www.unicode.org/versions/Unicode5.0.0/>
- [[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 <http://www.unicode.org/versions/Unicode5.1.0/>
- [[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), <http://www.unicode.org/versions/Unicode5.2.0/>
- [[Unicode6.0](#)] The Unicode Consortium, The Unicode Standard, Version 6.0.0 (Mountain View, CA: The Unicode Consortium, 2011. ISBN 978-1-936213-01-6) <http://www.unicode.org/versions/Unicode6.0.0/>
- [[Unicode6.1](#)] The Unicode Consortium, The Unicode Standard, Version 6.1.0 (Mountain View, CA: The Unicode Consortium, 2012. ISBN 978-1-936213-02-3) <http://www.unicode.org/versions/Unicode6.1.0/>
- [[Unicode6.2](#)] The Unicode Consortium, The Unicode Standard, Version 6.2.0

(Mountain View, CA: The Unicode Consortium, 2012.
ISBN 978-1-936213-07-8)
<http://www.unicode.org/versions/Unicode6.2.0/>

[[Unicode6.3](#)] The Unicode Consortium, The Unicode Standard,
Version 6.3.0
(Mountain View, CA: The Unicode Consortium, 2013.
ISBN 978-1-936213-08-5)
<http://www.unicode.org/versions/Unicode6.3.0/>

[[Unicode7.0](#)] The Unicode Consortium, The Unicode Standard,
Version 7.0.0
(Mountain View, CA: The Unicode Consortium, 2014.
ISBN 978-1-936213-09-2)
<http://www.unicode.org/versions/Unicode7.0.0/>

[[Unicode8.0](#)] The Unicode Consortium, The Unicode Standard,
Version 8.0.0
(Mountain View, CA: The Unicode Consortium, 2015.
ISBN 978-1-936213-10-8)
<http://www.unicode.org/versions/Unicode8.0.0/>

[[UTC](#)] Unicode Technical Committee
<http://www.unicode.org/consortium/utc.html>

[[UTN5](#)] Unicode Technical Note #5: Canonical Equivalence in
Applications
<http://www.unicode.org/notes/tn5>

[[UTR17](#)] Unicode Technical Report #17: Unicode Character
Encoding Model
Latest version:
<http://www.unicode.org/reports/tr17/>

[[UTR20](#)] Unicode Technical Report #20: Unicode in XML and
other Markup Languages
Latest version:
<http://www.unicode.org/reports/tr20/>

[[UTR23](#)] Unicode Technical Report #23: The Unicode Character
Property Model
Latest version:

<http://www.unicode.org/reports/tr23/>

[[UTR25](#)] Unicode Technical Report #25: Unicode Support for Mathematics

Latest version:

<http://www.unicode.org/reports/tr25/>

[[UTR33](#)] Unicode Technical Report #33: Unicode Conformance Model

Latest version:

<http://www.unicode.org/reports/tr33/>

[[UTR36](#)] Unicode Technical Report #36: Unicode Security Considerations

Latest version:

<http://www.unicode.org/reports/tr36/>

[[UTR50](#)] Unicode Technical Report #50: Unicode Vertical Text Layout

Latest version:

<http://www.unicode.org/reports/tr50/>

[[UTR51](#)] Unicode Technical Report #51: Unicode Emoji

Latest version:

<http://www.unicode.org/reports/tr51/>

[[UTS6](#)] Unicode Technical Standard #6: A Standard Compression Scheme for Unicode

Latest version:

<http://www.unicode.org/reports/tr6/>

Review Note: Prior to final release, links to the 8.0.0 versions of UTS #10 and UTS #46 resolve to the Proposed Updates of those UTSes, if they exist.

[[UTS10](#)] Unicode Technical Standard #10: Unicode Collation Algorithm

Latest version:

<http://www.unicode.org/reports/tr10/>

Version 8.0.0:

<http://www.unicode.org/reports/tr10/tr10-31.html>

- [[UTS18](#)] Unicode Technical Standard #18: Unicode Regular Expressions
Latest version:
<http://www.unicode.org/reports/tr18/>
- [[UTS22](#)] Unicode Technical Standard #22: Unicode Character Mapping Markup Language (CharMapML)
Latest version:
<http://www.unicode.org/reports/tr22/>
- [[UTS35](#)] Unicode Technical Standard #35: Unicode Locale Data Markup Language (LDML)
Latest version:
<http://www.unicode.org/reports/tr35/>
- [[UTS37](#)] Unicode Technical Standard #37: Unicode Ideographic Variation Database
Latest version:
<http://www.unicode.org/reports/tr37/>
- [[UTS39](#)] Unicode Technical Standard #39: Unicode Security Mechanisms
Latest version:
<http://www.unicode.org/reports/tr39/>
- [[UTS46](#)] Unicode Technical Standard #46: Unicode IDNA Compatibility Processing
Latest version:
<http://www.unicode.org/reports/tr46/>
Version 8.0.0:
<http://www.unicode.org/reports/tr46/tr46-14.html>
- [[Versions](#)] About Versions of the Unicode Standard
Information on version numbering, and citing and referencing the Unicode Standard, the Unicode Character Database, and Unicode Technical Reports:
<http://www.unicode.org/versions/>

2 References to Other Standards

- [[10646](#)] International Organization for Standardization,

ISO/IEC 10646:2014: Information Technology – Universal Coded Character Set (UCS), Fourth Edition
Available from the ISO/IEC ITTF website:

<http://standards.iso.org/ittf/PubliclyAvailableStandards/>

[HTML5]

Robin Berjon, et al., HTML 5: A Vocabulary and Associated APIs for HTML and XHTML, W3C Recommendation

<http://www.w3.org/TR/html5/>

[ISO15924]

International Organization for Standardization, ISO 15924:2004: Information and Documentation – Codes for the Representation of Names of Scripts
<http://www.unicode.org/iso15924/>

[ISO19757]

International Organization for Standardization, ISO/IEC 19757-2:2008: Information Technology – Document Schema Definition Language (DSDL) – Part 2: Regular-Grammar-Based Validation – RELAX NG, Second Edition

Available from the ISO/IEC ITTF website:

<http://standards.iso.org/ittf/PubliclyAvailableStandards/>

[ISO19757 Amd1]

International Organization for Standardization, ISO/IEC 19757-2:2003/Amd.1:2006: Information Technology – Document Schema Definition Language (DSDL) – Part 2: Regular-Grammar-Based Validation – RELAX NG – Amendment 1: Compact Syntax

Available from the ISO/IEC ITTF website:

<http://standards.iso.org/ittf/PubliclyAvailableStandards/>

[JIS]

Japanese Standards Association, JIS X 4051:2004: Formatting Rules for Japanese Documents 『日本語文書の組版方法』 Japanese Standards Association. 2004.

[KSX1026]

Korean Agency for Technology and Standards, KS X 1026-1:2007: Information Technology – Universal Multiple Octet Coded Character Set – Hangul – Part 1:

Hangul Processing Guide for Information Interchange
Korean Agency for Technology and Standards. 2008.

[XML]

Tim Bray, et al., Extensible Markup Language (XML)
1.0, Fifth Edition or later, W3C Recommendation
<http://www.w3.org/TR/xml/>

3 Other References

[Cedar97]

Cy Cedar, et al., 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]

Martin J. Dürst, Charlint – A Character Normalization Tool
<http://www.w3.org/International/charlint/>

[CharMatch]

Addison Phillips, Character Model for the World Wide Web:
String Matching and Searching, W3C Working Draft
<http://www.w3.org/TR/charmod-norm/>

[CharMod]

Martin J. Dürst, et al., François Yergeau, Richard Ishida,
Misha Wolf, and Tex Texin, W3C Character Model for the
World Wide Web 1.0: Fundamentals, W3C Recommendation
<http://www.w3.org/TR/charmod/>

[CharNorm]

François Yergeau, et al., Martin J. Dürst, Richard Ishida,
Misha Wolf, Tex Texin, and Addison Phillips, Character
Model for the World Wide Web 1.0: Normalization, W3C
Working Draft
<http://www.w3.org/TR/2012/WD-charmod-norm-20120501/>

[CharReq]

Martin J. Dürst, Requirements for String Identity Matching
and String Indexing, W3C Working Draft
<http://www.w3.org/TR/WD-charreq>

[Knuth78]

Donald E. Knuth, et al., and Michael F. Plass, Breaking
Lines into Paragraphs, republished in Digital Typography,
CSLI 78 (Stanford, California: CSLI 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, $T_{\epsilon}X$, the Program, Volume B of Computers
& Typesetting (Reading, MA: Addison–Wesley, 1986)

Copyright © 2001–2014 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.