

The Unicode Standard

Version 6.2 – Core Specification

To learn about the latest version of the Unicode Standard, see <http://www.unicode.org/versions/latest/>.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed with initial capital letters or in all capitals.

Unicode and the Unicode Logo are registered trademarks of Unicode, Inc., in the United States and other countries.

The authors and publisher have taken care in the preparation of this specification, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

The *Unicode Character Database* and other files are provided as-is by Unicode, Inc. No claims are made as to fitness for any particular purpose. No warranties of any kind are expressed or implied. The recipient agrees to determine applicability of information provided.

Copyright © 1991–2012 Unicode, Inc.

All rights reserved. This publication is protected by copyright, and permission must be obtained from the publisher prior to any prohibited reproduction. For information regarding permissions, inquire at <http://www.unicode.org/reporting.html>. For information about the Unicode terms of use, please see <http://www.unicode.org/copyright.html>.

The Unicode Standard / the Unicode Consortium ; edited by Julie D. Allen ... [et al.]. — Version 6.2.

Includes bibliographical references and index.

ISBN 978-1-936213-07-8 (<http://www.unicode.org/versions/Unicode6.2.0/>)

1. Unicode (Computer character set) I. Allen, Julie D. II. Unicode Consortium.

QA268.U545 2012

ISBN 978-1-936213-07-8

Published in Mountain View, CA

September 2012

Contents

Figures	xvii
Tables	xxi
Preface	xxv
Why Unicode?	xxv
What's New?	xxv
Organization of This Standard	xxv
Unicode Standard Annexes	xxvii
The Unicode Character Database	xxviii
Unicode Code Charts	xxviii
Unicode Technical Standards and Unicode Technical Reports	xxviii
Updates and Errata	xxviii
Acknowledgements	xxix
1 Introduction	1
1.1 Coverage	2
Standards Coverage	2
New Characters	3
1.2 Design Goals	3
1.3 Text Handling	4
Characters and Glyphs	5
Text Elements	5
2 General Structure	7
2.1 Architectural Context	7
Basic Text Processes	8
Text Elements, Characters, and Text Processes	8
Text Processes and Encoding	9
2.2 Unicode Design Principles	10
Universality	10
Efficiency	11
Characters, Not Glyphs	11
Semantics	14
Plain Text	14
Logical Order	15
Unification	17
Dynamic Composition	18
Stability	18
Convertibility	19
2.3 Compatibility Characters	19
Compatibility Variants	20
Compatibility Decomposable Characters	20
2.4 Code Points and Characters	21
Types of Code Points	22
2.5 Encoding Forms	24
UTF-32	26
UTF-16	27
UTF-8	27
Comparison of the Advantages of UTF-32, UTF-16, and UTF-8	28

2.6	Encoding Schemes	30
2.7	Unicode Strings	32
2.8	Unicode Allocation	33
	Planes	33
	Allocation Areas and Character Blocks	34
	Assignment of Code Points	35
2.9	Details of Allocation	35
	Plane 0 (BMP)	37
	Plane 1 (SMP)	38
	Plane 2 (SIP)	40
	Other Planes	40
2.10	Writing Direction	40
2.11	Combining Characters	41
	Sequence of Base Characters and Diacritics	42
	Multiple Combining Characters	43
	Ligated Multiple Base Characters	45
	Exhibiting Nonspacing Marks in Isolation	45
	“Characters” and Grapheme Clusters	46
2.12	Equivalent Sequences and Normalization	46
	Normalization	47
	Decompositions	48
	Non-decomposition of Overlaid Diacritics	49
2.13	Special Characters and Noncharacters	50
	Special Noncharacter Code Points	50
	Byte Order Mark (BOM)	50
	Layout and Format Control Characters	51
	The Replacement Character	51
	Control Codes	51
2.14	Conforming to the Unicode Standard	51
	Characteristics of Conformant Implementations	52
	Unacceptable Behavior	52
	Acceptable Behavior	53
	Supported Subsets	53
3	Conformance	55
3.1	Versions of the Unicode Standard	55
	Stability	56
	Version Numbering	56
	Errata and Corrigenda	57
	References to the Unicode Standard	57
	Precision in Version Citation	57
	References to Unicode Character Properties	58
	References to Unicode Algorithms	58
3.2	Conformance Requirements	59
	Code Points Unassigned to Abstract Characters	59
	Interpretation	59
	Modification	61
	Character Encoding Forms	62
	Character Encoding Schemes	62
	Bidirectional Text	63
	Normalization Forms	63
	Normative References	63
	Unicode Algorithms	63

	Default Casing Algorithms	64
	Unicode Standard Annexes	64
3.3	Semantics	65
	Definitions	65
	Character Identity and Semantics	65
3.4	Characters and Encoding	67
3.5	Properties	70
	Types of Properties	70
	Property Values	71
	Default Property Values	72
	Classification of Properties by Their Values	73
	Property Status	73
	Context Dependence	76
	Stability of Properties	76
	Simple and Derived Properties	78
	Property Aliases	78
	Private Use	79
3.6	Combination	79
	Combining Character Sequences	79
	Grapheme Clusters	81
	Application of Combining Marks	83
3.7	Decomposition	87
	Compatibility Decomposition	87
	Canonical Decomposition	88
3.8	Surrogates	89
3.9	Unicode Encoding Forms	89
	UTF-32	93
	UTF-16	93
	UTF-8	94
	Encoding Form Conversion	95
	Constraints on Conversion Processes	96
3.10	Unicode Encoding Schemes	98
3.11	Normalization Forms	101
	Normalization Stability	102
	Combining Classes	102
	Specification of Unicode Normalization Forms	103
	Starters	103
	Canonical Ordering Algorithm	104
	Canonical Composition Algorithm	105
	Definition of Normalization Forms	107
3.12	Conjoining Jamo Behavior	107
	Definitions	107
	Hangul Syllable Decomposition	109
	Hangul Syllable Composition	111
	Hangul Syllable Name Generation	112
	Sample Code for Hangul Algorithms	112
3.13	Default Case Algorithms	115
	Definitions	115
	Default Case Conversion	117
	Default Case Folding	117
	Default Case Detection	118
	Default Caseless Matching	119

4	Character Properties	121
4.1	Unicode Character Database	122
4.2	Case	124
	Definitions of Case and Casing	124
	Case Mapping	126
4.3	Combining Classes	126
	Reordrant, Split, and Subjoined Combining Marks	127
4.4	Directionality	130
4.5	General Category	130
4.6	Numeric Value	133
	Ideographic Numeric Values	133
4.7	Bidi Mirrored	135
4.8	Name	135
	Unicode Name Property	137
	Code Point Labels	138
	Use of Character Names in APIs and User Interfaces	138
4.9	Unicode 1.0 Names	139
4.10	Letters, Alphabetic, and Ideographic	139
4.11	Properties Related to Text Boundaries	140
4.12	Characters with Unusual Properties	140
5	Implementation Guidelines	145
5.1	Data Structures for Character Conversion	145
	Issues	145
	Multistage Tables	146
5.2	Programming Languages and Data Types	147
	Unicode Data Types for C	147
5.3	Unknown and Missing Characters	148
5.4	Handling Surrogate Pairs in UTF-16	149
5.5	Handling Numbers	151
5.6	Normalization	152
5.7	Compression	153
5.8	Newline Guidelines	153
	Definitions	153
	Line Separator and Paragraph Separator	154
	Recommendations	155
5.9	Regular Expressions	156
5.10	Language Information in Plain Text	157
	Requirements for Language Tagging	157
	Language Tags and Han Unification	157
5.11	Editing and Selection	158
	Consistent Text Elements	158
5.12	Strategies for Handling Nonspacing Marks	159
	Keyboard Input	160
	Truncation	161
5.13	Rendering Nonspacing Marks	162
	Canonical Equivalence	164
	Positioning Methods	165
5.14	Locating Text Element Boundaries	167
5.15	Identifiers	167

5.16	Sorting and Searching	167
	Culturally Expected Sorting and Searching	168
	Language-Insensitive Sorting	168
	Searching	168
	Sublinear Searching	169
5.17	Binary Order	170
	UTF-8 in UTF-16 Order	170
	UTF-16 in UTF-8 Order	171
5.18	Case Mappings	172
	Titlecasing	172
	Complications for Case Mapping	172
	Reversibility	174
	Caseless Matching	175
	Normalization and Casing	177
5.19	Mapping Compatibility Variants	177
5.20	Unicode Security	179
5.21	Ignoring Characters in Processing	180
	Characters Ignored in Text Segmentation	181
	Characters Ignored in Line Breaking	181
	Characters Ignored in Cursive Joining	181
	Characters Ignored in Identifiers	182
	Characters Ignored in Searching and Sorting	182
	Characters Ignored for Display	183
5.22	Best Practice for U+FFFD Substitution	185
6	Writing Systems and Punctuation	187
6.1	Writing Systems	188
6.2	General Punctuation	191
	Blocks Devoted to Punctuation	193
	Format Control Characters	193
	Space Characters	194
	Dashes and Hyphens	195
	Paired Punctuation	197
	Language-Based Usage of Quotation Marks	197
	Apostrophes	200
	Other Punctuation	200
	Archaic Punctuation and Editorial Marks	204
	Indic Punctuation	206
	CJK Punctuation	207
	Unknown or Unavailable Ideographs	208
	CJK Compatibility Forms	209
7	European Alphabetic Scripts	211
7.1	Latin	212
	Letters of Basic Latin: U+0041–U+007A	215
	Letters of the Latin-1 Supplement: U+00C0–U+00FF	215
	Latin Extended-A: U+0100–U+017F	216
	Latin Extended-B: U+0180–U+024F	216
	IPA Extensions: U+0250–U+02AF	218
	Phonetic Extensions: U+1D00–U+1DBF	219
	Latin Extended Additional: U+1E00–U+1EFF	220
	Latin Extended-C: U+2C60–U+2C7F	220
	Latin Extended-D: U+A720–U+A7FF	221
	Latin Ligatures: U+FB00–U+FB06	222

7.2	Greek	222
	Greek: U+0370–U+03FF	222
	Greek Extended: U+1F00–U+1FFF	225
	Ancient Greek Numbers: U+10140–U+1018F	226
7.3	Coptic	227
7.4	Cyrillic	229
	Cyrillic: U+0400–U+04FF	230
	Cyrillic Supplement: U+0500–U+052F	230
	Cyrillic Extended-A: U+2DE0–U+2DFF	231
	Cyrillic Extended-B: U+A640–U+A69F	231
7.5	Glagolitic	231
7.6	Armenian	232
7.7	Georgian	233
7.8	Modifier Letters	235
	Spacing Modifier Letters: U+02B0–U+02FF	236
	Modifier Tone Letters: U+A700–U+A71F	237
7.9	Combining Marks	238
	Combining Diacritical Marks: U+0300–U+036F	242
	Combining Diacritical Marks Supplement: U+1DC0–U+1DFF	242
	Combining Marks for Symbols: U+20D0–U+20FF	242
	Combining Half Marks: U+FE20–U+FE2F	243
	Combining Marks in Other Blocks	243
8	Middle Eastern Scripts	245
8.1	Hebrew	246
	Hebrew: U+0590–U+05FF	246
	Alphabetic Presentation Forms: U+FB1D–U+FB4F	250
8.2	Arabic	250
	Arabic: U+0600–U+06FF	250
	Arabic Cursive Joining	256
	Arabic Ligatures	258
	Arabic Joining Groups	259
	Arabic Supplement: U+0750–U+077F	265
	Arabic Extended-A: U+08A0–U+08FF	265
	Arabic Presentation Forms-A: U+FB50–U+FDFF	265
	Arabic Presentation Forms-B: U+FE70–U+FEFF	266
8.3	Syriac	266
	Syriac: U+0700–U+074F	266
	Syriac Shaping	269
8.4	Samaritan	273
8.5	Thaana	274
9	South Asian Scripts-I	277
9.1	Devanagari	278
	Devanagari: U+0900–U+097F	278
	Principles of the Devanagari Script	279
	Rendering Devanagari	284
	Devanagari Digits, Punctuation, and Symbols	291
	Extensions in the Main Devanagari Block	292
	Devanagari Extended: U+A8E0–U+A8FF	293
	Vedic Extensions: U+1CD0–U+1CFF	294
9.2	Bengali (Bangla)	295
9.3	Gurmukhi	300

9.4	Gujarati	303
9.5	Oriya	304
9.6	Tamil	306
	Tamil: U+0B80–U+0BFF	306
	Tamil Vowels	307
	Tamil Ligatures	308
	Tamil Named Character Sequences	311
9.7	Telugu	313
9.8	Kannada	315
	Kannada: U+0C80–U+0CFF	315
	Principles of the Kannada Script	315
	Rendering Kannada	317
9.9	Malayalam	317
10	South Asian Scripts-II	323
10.1	Sinhala	324
10.2	Tibetan	325
10.3	Lepcha	335
10.4	Phags-pa	336
10.5	Limbu	342
10.6	Syloti Nagri	344
10.7	Kaithi	345
10.8	Saurashtra	347
10.9	Sharada	348
10.10	Takri	349
10.11	Chakma	350
10.12	Meetei Mayek	351
10.13	Ol Chiki	353
10.14	Sora Sompeng	354
10.15	Kharoshthi	355
	Kharoshthi: U+10A00–U+10A5F	355
	Rendering Kharoshthi	356
10.16	Brahmi	359
11	Southeast Asian Scripts	363
11.1	Thai	364
11.2	Lao	366
11.3	Myanmar	368
	Myanmar: U+1000–U+109F	368
	Myanmar Extended-A: U+AA60–U+AA7F	371
	Khamti Shan	372
	Aiton and Phake	373
11.4	Khmer	374
	Khmer: U+1780–U+17FF	374
	Principles of the Khmer Script	374
	Khmer Symbols: U+19E0–U+19FF	383
11.5	Tai Le	383
11.6	New Tai Lue	384
11.7	Tai Tham	385
11.8	Tai Viet	387

11.9 Kayah Li	389
11.10 Cham	390
11.11 Philippine Scripts	392
Tagalog: U+1700–U+171F	392
Hanunóo: U+1720–U+173F	392
Buhid: U+1740–U+175F	392
Tagbanwa: U+1760–U+177F	392
Principles of the Philippine Scripts	392
11.12 Buginese	393
11.13 Balinese	394
11.14 Javanese	399
11.15 Rejang	401
11.16 Batak	402
11.17 Sundanese	403
12 East Asian Scripts	405
12.1 Han	406
CJK Unified Ideographs	406
CJK Standards	407
Blocks Containing Han Ideographs	409
General Characteristics of Han Ideographs	411
Principles of Han Unification	414
Unification Rules	415
Abstract Shape	416
Han Ideograph Arrangement	418
Radical-Stroke Indices	419
Mappings for Han Ideographs	419
CJK Unified Ideographs Extension B: U+20000–U+2A6D6	420
CJK Unified Ideographs Extension C: U+2A700–U+2B734	420
CJK Unified Ideographs Extension D: U+2B740–U+2B81D	420
CJK Compatibility Ideographs: U+F900–U+FAFF	420
CJK Compatibility Supplement: U+2F800–U+2FA1D	421
Kanbun: U+3190–U+319F	421
Symbols Derived from Han Ideographs	421
CJK and KangXi Radicals: U+2E80–U+2FD5	421
CJK Additions from HKSCS and GB 18030	422
CJK Strokes: U+31C0–U+31EF	423
12.2 Ideographic Description Characters	423
12.3 Bopomofo	426
12.4 Hiragana and Katakana	428
Hiragana: U+3040–U+309F	428
Katakana: U+30A0–U+30FF	428
Katakana Phonetic Extensions: U+31F0–U+31FF	429
Kana Supplement U+1B000–U+1B0FF	429
12.5 Halfwidth and Fullwidth Forms	429
12.6 Hangul	430
Hangul Jamo: U+1100–U+11FF	430
Hangul Jamo Extended-A: U+A960–U+A97F	431
Hangul Jamo Extended-B: U+D7B0–U+D7FF	431
Hangul Compatibility Jamo: U+3130–U+318F	431
Hangul Syllables: U+AC00–U+D7A3	432
12.7 Yi	433

13 Additional Modern Scripts	437
13.1 Ethiopic	438
Ethiopic: U+1200–U+137F	438
Ethiopic Extensions	440
13.2 Mongolian	440
13.3 Osmanya	447
13.4 Tifinagh	448
13.5 N’Ko	450
13.6 Vai	454
13.7 Bamum	455
Bamum: U+A6A0–U+A6FF	455
Bamum Supplement: U+16800–U+16A3F	456
13.8 Cherokee	456
13.9 Canadian Aboriginal Syllabics	457
Canadian Aboriginal Syllabics: U+1400–U+167F	457
Canadian Aboriginal Syllabics Extended: U+18B0–U+18FF	458
13.10 Deseret	459
13.11 Shavian	461
13.12 Lisu	461
13.13 Miao	463
14 Additional Ancient and Historic Scripts	465
14.1 Ogham	466
14.2 Old Italic	467
14.3 Runic	469
14.4 Gothic	471
14.5 Old Turkic	472
14.6 Linear B	473
Linear B Syllabary: U+10000–U+1007F	473
Linear B Ideograms: U+10080–U+100FF	473
Aegean Numbers: U+10100–U+1013F	473
14.7 Cypriot Syllabary	474
14.8 Ancient Anatolian Alphabets	474
Lycian: U+10280–U+1029F	474
Carian: U+102A0–U+102DF	474
Lydian: U+10920–U+1093F	474
14.9 Old South Arabian	475
14.10 Phoenician	477
14.11 Imperial Aramaic	478
14.12 Mandaic	479
14.13 Inscriptional Parthian and Inscriptional Pahlavi	481
14.14 Avestan	482
14.15 Ugaritic	483
14.16 Old Persian	484
14.17 Sumero-Akkadian	485
Cuneiform: U+12000–U+123FF	485
Cuneiform Numbers and Punctuation: U+12400–U+1247F	487
14.18 Egyptian Hieroglyphs	487
14.19 Meroitic Hieroglyphs and Meroitic Cursive	491

15 Symbols	493
15.1 Currency Symbols	494
15.2 Letterlike Symbols	496
Letterlike Symbols: U+2100–U+214F	496
Mathematical Alphanumeric Symbols: U+1D400–U+1D7FF	498
Mathematical Alphabets	498
Fonts Used for Mathematical Alphabets	500
Arabic Mathematical Alphabetic Symbols: U+1EE00–U+1EEFF	501
15.3 Numerals	502
Decimal Digits	502
Other Digits	504
Non-Decimal Radix Systems	505
Acrophonic Systems and Other Letter-based Numbers	506
Rumi Numeral Forms: U+10E60–U+10E7E	507
CJK Numerals	507
Fractions	509
Common Indic Number Forms: U+A830–U+A83F	509
15.4 Superscript and Subscript Symbols	510
Superscripts and Subscripts: U+2070–U+209F	510
15.5 Mathematical Symbols	511
Mathematical Operators: U+2200–U+22FF	512
Supplements to Mathematical Symbols and Arrows	513
Supplemental Mathematical Operators: U+2A00–U+2AFF	514
Miscellaneous Mathematical Symbols-A: U+27C0–U+27EF	514
Miscellaneous Mathematical Symbols-B: U+2980–U+29FF	514
Miscellaneous Symbols and Arrows: U+2B00–U+2B7F	515
Arrows: U+2190–U+21FF	515
Supplemental Arrows	515
Standardized Variants of Mathematical Symbols	516
15.6 Invisible Mathematical Operators	516
15.7 Technical Symbols	517
Control Pictures: U+2400–U+243F	517
Miscellaneous Technical: U+2300–U+23FF	517
Optical Character Recognition: U+2440–U+245F	520
15.8 Geometrical Symbols	520
Box Drawing and Block Elements	521
Geometric Shapes: U+25A0–U+25FF	521
15.9 Miscellaneous Symbols	522
Miscellaneous Symbols: U+2600–U+26FF	523
Miscellaneous Symbols and Pictographs: U+1F300–U+1F5FF	523
Emoticons: U+1F600–U+1F64F	525
Transport and Map Symbols: U+1F680–U+1F6FF	526
Dingbats: U+2700–U+27BF	526
Alchemical Symbols: U+1F700–U+1F77F	527
Mahjong Tiles: U+1F000–U+1F02F	527
Domino Tiles: U+1F030–U+1F09F	528
Playing Cards: U+1F0A0–U+1F0FF	528
Yijing Hexagram Symbols: U+4DC0–U+4DFF	529
Tai Xuan Jing Symbols: U+1D300–U+1D356	529
Ancient Symbols: U+10190–U+101CF	530
Phaistos Disc Symbols: U+101D0–U+101FF	530

15.10 Enclosed and Square	531
Enclosed Alphanumerics: U+2460–U+24FF	533
Enclosed CJK Letters and Months: U+3200–U+32FF	533
CJK Compatibility: U+3300–U+33FF	533
Enclosed Alphanumeric Supplement: U+1F100–U+1F1FF	534
Enclosed Ideographic Supplement: U+1F200–U+1F2FF	534
15.11 Braille	534
15.12 Western Musical Symbols	536
15.13 Byzantine Musical Symbols	540
15.14 Ancient Greek Musical Notation	540
16 Special Areas and Format Characters	543
16.1 Control Codes	544
Representing Control Sequences	544
Specification of Control Code Semantics	545
16.2 Layout Controls	545
Line and Word Breaking	546
Cursive Connection and Ligatures	548
Combining Grapheme Joiner	551
Bidirectional Ordering Controls	553
Stateful Format Controls	553
16.3 Deprecated Format Characters	554
16.4 Variation Selectors	556
16.5 Private-Use Characters	557
Private Use Area: U+E000–U+F8FF	558
Supplementary Private Use Areas	559
16.6 Surrogates Area	559
16.7 Noncharacters	560
16.8 Specials	561
Byte Order Mark (BOM): U+FEFF	561
Specials: U+FFF0–U+FFF8	563
Annotation Characters: U+FFF9–U+FFFB	563
Replacement Characters: U+FFFC–U+FFFD	564
16.9 Deprecated Tag Characters	565
Deprecated Tag Characters: U+E0000–U+E007F	565
Syntax for Embedding Tags	565
Working with Language Tags	567
Unicode Conformance Issues	568
Formal Tag Syntax	569
17 About the Code Charts	571
17.1 Character Names List	571
Images in the Code Charts and Character Lists	572
Special Characters and Code Points	573
Character Names	574
Informative Aliases	574
Normative Aliases	575
Cross References	575
Information About Languages	575
Case Mappings	576
Decompositions	576
Standardized Variation Sequences	577
Subheads	577

17.2	CJK Unified and Compatibility Ideographs	577
	CJK Unified Ideographs	578
	Compatibility Ideographs	579
17.3	Hangul Syllables	579
A	Notational Conventions	581
	Code Points	581
	Character Names	581
	Character Blocks	581
	Sequences	582
	Rendering	582
	Properties and Property Values	582
	Miscellaneous	582
	Extended BNF	583
	Operators	584
B	Unicode Publications and Resources	585
B.1	The Unicode Consortium	585
	The Unicode Technical Committee	585
	Other Activities	586
B.2	Unicode Publications	586
B.3	Unicode Technical Standards	586
	UTS #6: A Standard Compression Scheme for Unicode	586
	UTS #10: Unicode Collation Algorithm	587
	UTS #18: Unicode Regular Expressions	587
	UTS #22: Character Mapping Markup Language (CharMapML)	587
	UTS #35: Unicode Locale Data Markup Language (LDML)	587
	UTS #37: Unicode Ideographic Variation Database	587
	UTS #39: Unicode Security Mechanisms	587
B.4	Unicode Technical Reports	587
	UTR #16: UTF-EBCDIC	587
	UTR #17: Unicode Character Encoding Model	588
	UTR #20: Unicode in XML and Other Markup Languages	588
	UTR #23: The Unicode Character Property Model	588
	UTR #25: Unicode Support for Mathematics	588
	UTR #26: Compatibility Encoding Scheme for UTF-16: 8-Bit (CESU-8) ..	588
	UTR #33: Unicode Conformance Model	588
	UTR #36: Unicode Security Considerations	588
B.5	Unicode Technical Notes	589
B.6	Other Unicode Online Resources	589
	Unicode Online Resources	589
	How to Contact the Unicode Consortium	591
C	Relationship to ISO/IEC 10646	593
C.1	History	593
C.2	Encoding Forms in ISO/IEC 10646	597
	Zero Extending	597
C.3	UTF-8 and UTF-16	598
	UTF-8	598
	UTF-16	598
C.4	Synchronization of the Standards	598
C.5	Identification of Features for the Unicode Standard	598
C.6	Character Names	599
C.7	Character Functional Specifications	599

D	Changes from Previous Versions	601
D.1	Versions of the Unicode Standard	601
D.2	Clause and Definition Updates	603
E	Han Unification History	605
E.1	Development of the URO	605
E.2	Ideographic Rapporteur Group	606
F	Documentation of CJK Strokes	609
R	References	615
R.1	Source Standards and Specifications	615
R.2	Source Dictionaries for Han Unification	622
R.3	Other Sources for the Unicode Standard	622
R.4	Selected Resources: Technical	638
R.5	Selected Resources: Other	640
I	General Index	645

