

Chapter 4

Character Properties

This chapter describes the attributes of character properties defined by the Unicode Standard and gives mappings of characters to specific character properties. As specified in *Chapter 3, Conformance*, the Unicode Standard defines both normative and informative properties. The normative character properties of the Unicode Standard are given in the Table 4-1, and the informative character properties are given in Table 4-2. Also included in the tables is the location of each property's description and where the list of characters with that property can be found.

Table 4-1. Normative Character Properties

Property	Description	List
Case	Chapter 4	Chapter 7 ^a
Combining Classes	Chapter 3	Chapter 4
Combining Jamo	Chapter 3	Chapter 7
Decomposition (Canonical and Compatibility)	Chapter 3	Chapter 7
Directionality	Chapter 4	Chapter 4
Jamo Short Name	Chapter 3	Chapter 4
Numeric Value	Chapter 4	Chapter 4
Private Use	Chapter 3	None
Special Character Properties	Chapter 6	Chapter 3
Surrogate Property	Chapter 3	None
Mirrored	Chapter 3	Chapter 4
Unicode 2.0 Character Names	Chapter 7	Chapter 7

a. These properties are listed by individual character where they are not obvious. Full listings are provided in the *Unicode Character Database* on the CD-ROM.

Table 4-2. Informative Character Properties

Property	Description	List
Case Mapping	Chapter 4	CD-ROM
Dashes	Chapter 6	Chapter 6
Fullwidth and Halfwidth	Chapter 6	Chapter 7
Letters (Alphabetic and Ideographic)	Chapter 4	CD-ROM
Line Breaking	Chapter 6	Chapter 3
Mathematical Property	Chapter 4	Chapter 4
Spaces	Chapter 6	Chapter 6
Unicode 1.0 Names	Chapter 4	CD-ROM

Disclaimer

The content of all character property tables has been verified as far as possible by the Unicode Consortium. However, the Unicode Consortium does not guarantee that the tables

printed in this volume or on the CD-ROM are correct in every detail, and it is not responsible for errors that may occur either in the character property tables or in software that implements these tables.

4.1 Case

Case is a normative property of characters in certain alphabets whereby characters are considered to be variants of a single letter. These variants, which may differ markedly in shape and size, are called the *uppercase* letter (also known as *capital* or *majuscule*) and the *lowercase* letter (also known as *small* or *minuscule*). The uppercase letter is generally larger than the lowercase letter.

Because of the inclusion of certain composite characters for compatibility, such as U+01F1 LATIN CAPITAL LETTER DZ, there is a third case, called *titlecase*, which is used where the first character of a word is to be capitalized. An example of such a character is U+01F2 LATIN CAPITAL LETTER D WITH SMALL LETTER Z. The three case forms are UPPERCASE, Titlecase, lowercase.

For those scripts that have case (Latin, Greek, Cyrillic, and Armenian), the case of a Unicode character can be obtained from the character's name. This is true for only these four scripts. Uppercase characters contain the word *capital* in their names. Lowercase characters contain the word *small*. The word *small* in the names of characters from scripts other than the four just listed has nothing to do with case. (Note that while the archaic Georgian script contained upper- and lowercase pairs, they are rarely used in modern Georgian. See the Georgian subsection in *Section 6.1, General Scripts Area*).

Case mappings. The lowercase letter default case mapping is between the small character and the capital character. The Unicode Standard case mapping tables, which are informative, are on the CD-ROM.

In a few instances, upper- and lowercase mappings may differ from language to language between writing systems that employ the same letters. Examples include Turkish ("ı", U+0131 LATIN SMALL LETTER DOTLESS I maps to "I", U+0049 LATIN CAPITAL LETTER I) and French ("é", U+00E9 LATIN SMALL LETTER E WITH ACUTE generally maps to "É", U+00C9 LATIN CAPITAL LETTER E WITH ACUTE, but in some circumstances may map to "E", U+0045 LATIN CAPITAL LETTER E). However, in general the vast majority of case mappings are uniform across languages.

It is important to note that casing operations do not always provide a round-trip mapping. Also, since many characters are really caseless (most of the IPA block, for example), uppercasing a string does not mean that it will no longer contain any lowercase letters.

Because there are many more lowercase forms than there are uppercase or titlecase, it is recommended that the lowercase form be used for normalization, such as when strings are case-folded for loose comparison or indexing.

- ➔ Case correspondences are not always one-to-one: the result of case folding may be a different character length than in the source string (for example, U+00DF LATIN SMALL LETTER SHARP S (ß) becomes "SS" in uppercase).

4.2 Combining Classes

Each combining character has a normative canonical *combining class*. This class is used with the Canonical Ordering Algorithm to determine which combining characters interact typographically and to determine the canonical ordering of sequences of combining char-

acters. Class zero combining characters act like base letters for the purpose of determining canonical order. Combining characters with non-zero classes participate in re-ordering for the purpose of determining the canonical form of sequences of characters. (See Section 3.9, *Canonical Ordering Behavior*, for a description of the algorithm.)

The list of combining characters and their canonical combining class appears in Table 4-3. Most combining characters are non-spacing. The spacing, class zero combining characters are so noted.

Table 4-3. Combining Classes

Code	Class	Name
<i>Spacing</i>		
U+0903	0	DEVANAGARI SIGN VISARGA
U+093E	0	DEVANAGARI VOWEL SIGN AA
U+0940	0	DEVANAGARI VOWEL SIGN II
U+0949	0	DEVANAGARI VOWEL SIGN CANDRA O
U+094A	0	DEVANAGARI VOWEL SIGN SHORT O
U+094B	0	DEVANAGARI VOWEL SIGN O
U+094C	0	DEVANAGARI VOWEL SIGN AU
U+0982	0	BENGALI SIGN ANUSVARA
U+0983	0	BENGALI SIGN VISARGA
U+09BE	0	BENGALI VOWEL SIGN AA
U+09C0	0	BENGALI VOWEL SIGN II
U+09D7	0	BENGALI AU LENGTH MARK
U+0A3E	0	GURMUKHI VOWEL SIGN AA
U+0A40	0	GURMUKHI VOWEL SIGN II
U+0A83	0	GUJARATI SIGN VISARGA
U+0ABE	0	GUJARATI VOWEL SIGN AA
U+0AC0	0	GUJARATI VOWEL SIGN II
U+0AC9	0	GUJARATI VOWEL SIGN CANDRA O
U+0ACB	0	GUJARATI VOWEL SIGN O
U+0ACC	0	GUJARATI VOWEL SIGN AU
U+0B02	0	ORIYA SIGN ANUSVARA
U+0B03	0	ORIYA SIGN VISARGA
U+0B3E	0	ORIYA VOWEL SIGN AA
U+0B40	0	ORIYA VOWEL SIGN II
U+0B57	0	ORIYA AU LENGTH MARK
U+0B82	0	TAMIL SIGN ANUSVARA
U+0B83	0	TAMIL SIGN VISARGA
U+0BBE	0	TAMIL VOWEL SIGN AA
U+0BBF	0	TAMIL VOWEL SIGN I
U+0BC1	0	TAMIL VOWEL SIGN U
U+0BC2	0	TAMIL VOWEL SIGN UU
U+0BD7	0	TAMIL AU LENGTH MARK
U+0C01	0	TELUGU SIGN CANDRABINDU
U+0C02	0	TELUGU SIGN ANUSVARA
U+0C03	0	TELUGU SIGN VISARGA
U+0C41	0	TELUGU VOWEL SIGN U
U+0C42	0	TELUGU VOWEL SIGN UU
U+0C43	0	TELUGU VOWEL SIGN VOCALIC R
U+0C44	0	TELUGU VOWEL SIGN VOCALIC RR
U+0C82	0	KANNADA SIGN ANUSVARA
U+0C83	0	KANNADA SIGN VISARGA
U+0CBE	0	KANNADA VOWEL SIGN AA
U+0CC1	0	KANNADA VOWEL SIGN U
U+0CC2	0	KANNADA VOWEL SIGN UU
U+0CC3	0	KANNADA VOWEL SIGN VOCALIC R
U+0CC4	0	KANNADA VOWEL SIGN VOCALIC RR
U+0CD5	0	KANNADA LENGTH MARK
U+0CD6	0	KANNADA AI LENGTH MARK
U+0D02	0	MALAYALAM SIGN ANUSVARA
U+0D03	0	MALAYALAM SIGN VISARGA

Table 4-3. Combining Classes (Continued)

Code	Class	Name
U+0D3E	0	MALAYALAM VOWEL SIGN AA
U+0D3F	0	MALAYALAM VOWEL SIGN I
U+0D40	0	MALAYALAM VOWEL SIGN II
U+0D57	0	MALAYALAM AU LENGTH MARK
U+0F7F	0	TIBETAN SIGN RNAM BCAD
<i>Split</i>		
U+09CB	0	BENGALI VOWEL SIGN O
U+09CC	0	BENGALI VOWEL SIGN AU
U+0B48	0	ORIYA VOWEL SIGN AI
U+0B4B	0	ORIYA VOWEL SIGN O
U+0B4C	0	ORIYA VOWEL SIGN AU
U+0BCA	0	TAMIL VOWEL SIGN O
U+0BCB	0	TAMIL VOWEL SIGN OO
U+0BCC	0	TAMIL VOWEL SIGN AU
U+0CC0	0	KANNADA VOWEL SIGN II
U+0CC7	0	KANNADA VOWEL SIGN EE
U+0CC8	0	KANNADA VOWEL SIGN AI
U+0CCA	0	KANNADA VOWEL SIGN O
U+0CCB	0	KANNADA VOWEL SIGN OO
U+0D4A	0	MALAYALAM VOWEL SIGN O
U+0D4B	0	MALAYALAM VOWEL SIGN OO
U+0D4C	0	MALAYALAM VOWEL SIGN AU
<i>Reordrant</i>		
U+093F	0	DEVANAGARI VOWEL SIGN I
U+09BF	0	BENGALI VOWEL SIGN I
U+09C8	0	BENGALI VOWEL SIGN AI
U+09C7	0	BENGALI VOWEL SIGN E
U+0A3F	0	GURMUKHI VOWEL SIGN I
U+0ABF	0	GUJARATI VOWEL SIGN I
U+0B47	0	ORIYA VOWEL SIGN E
U+0BC6	0	TAMIL VOWEL SIGN E
U+0BC7	0	TAMIL VOWEL SIGN EE
U+0BC8	0	TAMIL VOWEL SIGN AI
U+0D46	0	MALAYALAM VOWEL SIGN E
U+0D47	0	MALAYALAM VOWEL SIGN EE
U+0D48	0	MALAYALAM VOWEL SIGN AI
<i>Enclosing</i>		
U+20DD	0	COMBINING ENCLOSING CIRCLE
U+20DE	0	COMBINING ENCLOSING SQUARE
U+20DF	0	COMBINING ENCLOSING DIAMOND
U+20E0	0	COMBINING ENCLOSING CIRCLE BACKSLASH
<i>Surrounding</i>		
U+06DD	0	ARABIC END OF AYAH
U+06DE	0	ARABIC START OF RUB EL HIZB
<i>Overlays /Interior</i>		
U+0334	1	COMBINING TILDE OVERLAY
U+0335	1	COMBINING SHORT STROKE OVERLAY
U+0336	1	COMBINING LONG STROKE OVERLAY
U+0337	1	COMBINING SHORT SOLIDUS OVERLAY
U+0338	1	COMBINING LONG SOLIDUS OVERLAY
U+20D2	1	COMBINING LONG VERTICAL LINE OVERLAY
U+20D3	1	COMBINING SHORT VERTICAL LINE OVERLAY
U+20D8	1	COMBINING RING OVERLAY
U+20D9	1	COMBINING CLOCKWISE RING OVERLAY
U+20DA	1	COMBINING ANTICLOCKWISE RING OVERLAY
U+20E1	1	COMBINING LEFT RIGHT ARROW ABOVE
<i>Tibetan Subjoined Letters</i>		
U+0F90	6	TIBETAN SUBJOINED LETTER KA
U+0F91	6	TIBETAN SUBJOINED LETTER KHA
U+0F92	6	TIBETAN SUBJOINED LETTER GA
U+0F93	6	TIBETAN SUBJOINED LETTER GHA

Table 4-3. Combining Classes (Continued)

Code	Class	Name
U+0F94	6	TIBETAN SUBJOINED LETTER NGA
U+0F95	6	TIBETAN SUBJOINED LETTER CA
U+0F97	6	TIBETAN SUBJOINED LETTER JA
U+0F99	6	TIBETAN SUBJOINED LETTER NYA
U+0F9A	6	TIBETAN SUBJOINED LETTER TTA
U+0F9B	6	TIBETAN SUBJOINED LETTER TTHA
U+0F9C	6	TIBETAN SUBJOINED LETTER DDA
U+0F9D	6	TIBETAN SUBJOINED LETTER DDHA
U+0F9E	6	TIBETAN SUBJOINED LETTER NNA
U+0F9F	6	TIBETAN SUBJOINED LETTER TA
U+0FA0	6	TIBETAN SUBJOINED LETTER THA
U+0FA1	6	TIBETAN SUBJOINED LETTER DA
U+0FA2	6	TIBETAN SUBJOINED LETTER DHA
U+0FA3	6	TIBETAN SUBJOINED LETTER NA
U+0FA4	6	TIBETAN SUBJOINED LETTER PA
U+0FA5	6	TIBETAN SUBJOINED LETTER PHA
U+0FA6	6	TIBETAN SUBJOINED LETTER BA
U+0FA7	6	TIBETAN SUBJOINED LETTER BHA
U+0FA8	6	TIBETAN SUBJOINED LETTER MA
U+0FA9	6	TIBETAN SUBJOINED LETTER TSA
U+0FAA	6	TIBETAN SUBJOINED LETTER TSHA
U+0FAB	6	TIBETAN SUBJOINED LETTER DZA
U+0FAC	6	TIBETAN SUBJOINED LETTER DZHA
U+0FAD	6	TIBETAN SUBJOINED LETTER WA
U+0FB1	6	TIBETAN SUBJOINED LETTER YA
U+0FB2	6	TIBETAN SUBJOINED LETTER RA
U+0FB3	6	TIBETAN SUBJOINED LETTER LA
U+0FB4	6	TIBETAN SUBJOINED LETTER SHA
U+0FB5	6	TIBETAN SUBJOINED LETTER SSA
U+0FB6	6	TIBETAN SUBJOINED LETTER SA
U+0FB7	6	TIBETAN SUBJOINED LETTER HA
U+0FB9	6	TIBETAN SUBJOINED LETTER KSSA
<i>Nuktas</i>		
U+093C	7	DEVANAGARI SIGN NUKTA
U+09BC	7	BENGALI SIGN NUKTA
U+0A3C	7	GURMUKHI SIGN NUKTA
U+0ABC	7	GUJARATI SIGN NUKTA
U+0B3C	7	ORIYA SIGN NUKTA
<i>Kana Voiced Marks</i>		
U+3099	8	COMBINING KATAKANA-HIRAGANA VOICED SOUND MARK
U+309A	8	COMBINING KATAKANA-HIRAGANA SEMI-VOICED SOUND MARK
<i>Viramas</i>		
U+094D	9	DEVANAGARI SIGN VIRAMA
U+09CD	9	BENGALI SIGN VIRAMA
U+0A4D	9	GURMUKHI SIGN VIRAMA
U+0ACD	9	GUJARATI SIGN VIRAMA
U+0B4D	9	ORIYA SIGN VIRAMA
U+0BCD	9	TAMIL SIGN VIRAMA
U+0C4D	9	TELUGU SIGN VIRAMA
U+0CCD	9	KANNADA SIGN VIRAMA
U+0D4D	9	MALAYALAM SIGN VIRAMA
U+0F84	9	TIBETAN MARK HALANTA
<i>Fixed Position Classes</i>		
U+05B0	10	HEBREW POINT SHEVA
U+05B1	11	HEBREW POINT HATAF SEGOL
U+05B2	12	HEBREW POINT HATAF PATAH
U+05B3	13	HEBREW POINT HATAF QAMATS
U+05B4	14	HEBREW POINT HIRIQ
U+05B5	15	HEBREW POINT TSERE
U+05B6	16	HEBREW POINT SEGOL
U+05B7	17	HEBREW POINT PATAH

Table 4-3. Combining Classes (Continued)

Code	Class	Name
U+05B8	18	HEBREW POINT QAMATS
U+05B9	19	HEBREW POINT HOLAM
U+05BB	20	HEBREW POINT QUBUTS
U+05BC	21	HEBREW POINT DAGESH OR MAPIQ
U+05BD	22	HEBREW POINT METEG
U+05BF	23	HEBREW POINT RAFE
U+05C1	24	HEBREW POINT SHIN DOT
U+05C2	25	HEBREW POINT SIN DOT
U+FB1E	26	HEBREW POINT JUDEO-SPANISH VARIKA
U+064B	27	ARABIC FATHATAN
U+064C	28	ARABIC DAMMATAN
U+064D	29	ARABIC KASRATAN
U+064E	30	ARABIC FATHA
U+064F	31	ARABIC DAMMA
U+0650	32	ARABIC KASRA
U+0651	33	ARABIC SHADDA
U+0652	34	ARABIC SUKUN
U+0670	35	ARABIC LETTER SUPERSCRIPT ALEF
U+0902	36	DEVANAGARI SIGN ANUSVARA
U+0901	37	DEVANAGARI SIGN CANDRABINDU
U+0941	38	DEVANAGARI VOWEL SIGN U
U+0942	39	DEVANAGARI VOWEL SIGN UU
U+0943	40	DEVANAGARI VOWEL SIGN VOCALIC R
U+0944	41	DEVANAGARI VOWEL SIGN VOCALIC RR
U+0945	42	DEVANAGARI VOWEL SIGN CANDRA E
U+0946	43	DEVANAGARI VOWEL SIGN SHORT E
U+0947	44	DEVANAGARI VOWEL SIGN E
U+0948	45	DEVANAGARI VOWEL SIGN AI
U+0951	46	DEVANAGARI STRESS SIGN UDATTA
U+0952	47	DEVANAGARI STRESS SIGN ANUDATTA
U+0962	48	DEVANAGARI VOWEL SIGN VOCALIC L
U+0963	49	DEVANAGARI VOWEL SIGN VOCALIC LL
U+0981	50	BENGALI SIGN CANDRABINDU
U+09C1	51	BENGALI VOWEL SIGN U
U+09C2	52	BENGALI VOWEL SIGN UU
U+09C3	53	BENGALI VOWEL SIGN VOCALIC R
U+09C4	54	BENGALI VOWEL SIGN VOCALIC RR
U+09E2	55	BENGALI VOWEL SIGN VOCALIC L
U+09E3	56	BENGALI VOWEL SIGN VOCALIC LL
U+0A02	57	GURMUKHI SIGN BINDI
U+0A41	58	GURMUKHI VOWEL SIGN U
U+0A42	59	GURMUKHI VOWEL SIGN UU
U+0A47	60	GURMUKHI VOWEL SIGN EE
U+0A48	61	GURMUKHI VOWEL SIGN AI
U+0A4B	62	GURMUKHI VOWEL SIGN OO
U+0A4C	63	GURMUKHI VOWEL SIGN AU
U+0A70	64	GURMUKHI TIPPI
U+0A71	65	GURMUKHI ADDAK
U+0A82	66	GUJARATI SIGN ANUSVARA
U+0A81	67	GUJARATI SIGN CANDRABINDU
U+0AC1	68	GUJARATI VOWEL SIGN U
U+0AC2	69	GUJARATI VOWEL SIGN UU
U+0AC3	70	GUJARATI VOWEL SIGN VOCALIC R
U+0AC4	71	GUJARATI VOWEL SIGN VOCALIC RR
U+0AC5	72	GUJARATI VOWEL SIGN CANDRA E
U+0AC7	73	GUJARATI VOWEL SIGN E
U+0AC8	74	GUJARATI VOWEL SIGN AI
U+0B01	75	ORIYA SIGN CANDRABINDU
U+0B3F	76	ORIYA VOWEL SIGN I
U+0B41	77	ORIYA VOWEL SIGN U
U+0B42	78	ORIYA VOWEL SIGN UU

Table 4-3. Combining Classes (Continued)

Code	Class	Name
U+0B43	79	ORIYA VOWEL SIGN VOCALIC R
U+0BC0	80	TAMIL VOWEL SIGN II
U+0C3E	81	TELUGU VOWEL SIGN AA
U+0C3F	82	TELUGU VOWEL SIGN I
U+0C40	83	TELUGU VOWEL SIGN II
U+0C46	84	TELUGU VOWEL SIGN E
U+0C47	85	TELUGU VOWEL SIGN EE
U+0C48	86	TELUGU VOWEL SIGN AI
U+0C4A	87	TELUGU VOWEL SIGN O
U+0C4B	88	TELUGU VOWEL SIGN OO
U+0C4C	89	TELUGU VOWEL SIGN AU
U+0C55	90	TELUGU LENGTH MARK
U+0C56	91	TELUGU AI LENGTH MARK
U+0CBF	92	KANNADA VOWEL SIGN I
U+0CC6	93	KANNADA VOWEL SIGN E
U+0CCC	94	KANNADA VOWEL SIGN AU
U+0D41	95	MALAYALAM VOWEL SIGN U
U+0D42	96	MALAYALAM VOWEL SIGN UU
U+0D43	97	MALAYALAM VOWEL SIGN VOCALIC R
U+0E31	98	THAI CHARACTER MAI HAN-AKAT
U+0E34	99	THAI CHARACTER SARA I
U+0E35	100	THAI CHARACTER SARA II
U+0E36	101	THAI CHARACTER SARA UE
U+0E37	102	THAI CHARACTER SARA UEE
U+0E38	103	THAI CHARACTER SARA U
U+0E39	104	THAI CHARACTER SARA UU
U+0E3A	105	THAI CHARACTER PHINTHU
U+0E47	106	THAI CHARACTER MAITAIKHU
U+0E48	107	THAI CHARACTER MAI EK
U+0E49	108	THAI CHARACTER MAI THO
U+0E4A	109	THAI CHARACTER MAI TRI
U+0E4B	110	THAI CHARACTER MAI CHATTAWA
U+0E4C	111	THAI CHARACTER THANTHAKHAT
U+0E4D	112	THAI CHARACTER NIKHAHIT
U+0EB1	113	LAO VOWEL SIGN MAI KAN
U+0EB4	114	LAO VOWEL SIGN I
U+0EB5	115	LAO VOWEL SIGN II
U+0EB6	116	LAO VOWEL SIGN Y
U+0EB7	117	LAO VOWEL SIGN YY
U+0EB8	118	LAO VOWEL SIGN U
U+0EB9	119	LAO VOWEL SIGN UU
U+0EBB	120	LAO VOWEL SIGN MAI KON
U+0EBC	121	LAO SEMIVOWEL SIGN LO
U+0EC8	122	LAO TONE MAI EK
U+0EC9	123	LAO TONE MAI THO
U+0ECA	124	LAO TONE MAI TI
U+0ECB	125	LAO TONE MAI CATAWA
U+0ECC	126	LAO CANCELLATION MARK
U+0ECD	127	LAO NIGGAHITA
U+0E4E	128	THAI CHARACTER YAMAKKAN
U+0F71	129	TIBETAN VOWEL SIGN AA
U+0F72	130	TIBETAN VOWEL SIGN I
U+0F73	131	TIBETAN VOWEL SIGN II
U+0F74	132	TIBETAN VOWEL SIGN U
U+0F75	133	TIBETAN VOWEL SIGN UU
U+0F76	134	TIBETAN VOWEL SIGN VOCALIC R
U+0F77	135	TIBETAN VOWEL SIGN VOCALIC RR
U+0F78	136	TIBETAN VOWEL SIGN VOCALIC L
U+0F79	137	TIBETAN VOWEL SIGN VOCALIC LL
U+0F7A	138	TIBETAN VOWEL SIGN E
U+0F7B	139	TIBETAN VOWEL SIGN EE

Table 4-3. Combining Classes (Continued)

Code	Class	Name
U+0F7C	140	TIBETAN VOWEL SIGN O
U+0F7D	141	TIBETAN VOWEL SIGN OO
U+0F7E	142	TIBETAN VOWEL SIGN RJES SU NGA RO
U+0F80	143	TIBETAN VOWEL SIGN REVERSED I
U+0F81	144	TIBETAN VOWEL SIGN REVERSED II
<i>Below Attached</i>		
U+0321	202	COMBINING PALATALIZED HOOK BELOW
U+0322	202	COMBINING RETROFLEX HOOK BELOW
U+0327	202	COMBINING CEDILLA
U+0328	202	COMBINING OGONEK
<i>Above Right Attached</i>		
U+031B	216	COMBINING HORN
U+0F39	216	TIBETAN MARK TSA -PHRU
<i>Below Left</i>		
U+302A	218	IDEOGRAPHIC LEVEL TONE MARK
<i>Below</i>		
U+0316	220	COMBINING GRAVE ACCENT BELOW
U+0317	220	COMBINING ACUTE ACCENT BELOW
U+0318	220	COMBINING LEFT TACK BELOW
U+0319	220	COMBINING RIGHT TACK BELOW
U+031C	220	COMBINING LEFT HALF RING BELOW
U+031D	220	COMBINING UP TACK BELOW
U+031E	220	COMBINING DOWN TACK BELOW
U+031F	220	COMBINING PLUS SIGN BELOW
U+0320	220	COMBINING MINUS SIGN BELOW
U+0323	220	COMBINING DOT BELOW
U+0324	220	COMBINING DIAERESIS BELOW
U+0325	220	COMBINING RING BELOW
U+0326	220	COMBINING COMMA BELOW
U+0329	220	COMBINING VERTICAL LINE BELOW
U+032A	220	COMBINING BRIDGE BELOW
U+032B	220	COMBINING INVERTED DOUBLE ARCH BELOW
U+032C	220	COMBINING CARON BELOW
U+032D	220	COMBINING CIRCUMFLEX ACCENT BELOW
U+032E	220	COMBINING BREVE BELOW
U+032F	220	COMBINING INVERTED BREVE BELOW
U+0330	220	COMBINING TILDE BELOW
U+0331	220	COMBINING MACRON BELOW
U+0332	220	COMBINING LOW LINE
U+0333	220	COMBINING DOUBLE LOW LINE
U+0339	220	COMBINING RIGHT HALF RING BELOW
U+033A	220	COMBINING INVERTED BRIDGE BELOW
U+033B	220	COMBINING SQUARE BELOW
U+033C	220	COMBINING SEAGULL BELOW
U+0345	220	COMBINING GREEK YPOGEGRAMMENI
U+0591	220	HEBREW ACCENT ETNAHTA
U+0596	220	HEBREW ACCENT TIPEHA
U+059B	220	HEBREW ACCENT TEVIR
U+05A3	220	HEBREW ACCENT MUNAH
U+05A4	220	HEBREW ACCENT MAHAPAKH
U+05A5	220	HEBREW ACCENT MERKHA
U+05A6	220	HEBREW ACCENT MERKHA KEFULA
U+05A7	220	HEBREW ACCENT DARGA
U+05AA	220	HEBREW ACCENT YERAH BEN YOMO
U+06E3	220	ARABIC SMALL LOW SEEN
U+06EA	220	ARABIC EMPTY CENTRE LOW STOP
U+06ED	220	ARABIC SMALL LOW MEEM
U+0F18	220	TIBETAN ASTROLOGICAL SIGN -KHYUD PA
U+0F19	220	TIBETAN ASTROLOGICAL SIGN SDONG TSHUGS
U+0F35	220	TIBETAN MARK NGAS BZUNG NYI ZLA
U+0F37	220	TIBETAN MARK NGAS BZUNG SGOR RTAGS

Table 4-3. Combining Classes (Continued)

Code	Class	Name
U+0F3E	220	TIBETAN SIGN YAR TSHES
U+0F3F	220	TIBETAN SIGN MAR TSHES
<i>Below Right</i>		
U+059A	222	HEBREW ACCENT YETIV
U+05AD	222	HEBREW ACCENT DEHI
U+302D	222	IDEOGRAPHIC ENTERING TONE MARK
<i>Left</i>		
U+302E	224	HANGUL SINGLE DOT TONE MARK
U+302F	224	HANGUL DOUBLE DOT TONE MARK
<i>Above Left</i>		
U+302B	228	IDEOGRAPHIC RISING TONE MARK
<i>Above</i>		
U+0300	230	COMBINING GRAVE ACCENT
U+0301	230	COMBINING ACUTE ACCENT
U+0302	230	COMBINING CIRCUMFLEX ACCENT
U+0303	230	COMBINING TILDE
U+0304	230	COMBINING MACRON
U+0305	230	COMBINING OVERLINE
U+0306	230	COMBINING BREVE
U+0307	230	COMBINING DOT ABOVE
U+0308	230	COMBINING DIAERESIS
U+0309	230	COMBINING HOOK ABOVE
U+030A	230	COMBINING RING ABOVE
U+030B	230	COMBINING DOUBLE ACUTE ACCENT
U+030C	230	COMBINING CARON
U+030D	230	COMBINING VERTICAL LINE ABOVE
U+030E	230	COMBINING DOUBLE VERTICAL LINE ABOVE
U+030F	230	COMBINING DOUBLE GRAVE ACCENT
U+0310	230	COMBINING CANDRABINDU
U+0311	230	COMBINING INVERTED BREVE
U+0312	230	COMBINING TURNED COMMA ABOVE
U+0313	230	COMBINING COMMA ABOVE
U+0314	230	COMBINING REVERSED COMMA ABOVE
U+033D	230	COMBINING X ABOVE
U+033E	230	COMBINING VERTICAL TILDE
U+033F	230	COMBINING DOUBLE OVERLINE
U+0340	230	COMBINING GRAVE TONE MARK
U+0341	230	COMBINING ACUTE TONE MARK
U+0342	230	COMBINING GREEK PERISPOMENI
U+0343	230	COMBINING GREEK KORONIS
U+0344	230	COMBINING GREEK DIALYTIKA TONOS
U+0483	230	COMBINING CYRILLIC TITLO
U+0484	230	COMBINING CYRILLIC PALATALIZATION
U+0485	230	COMBINING CYRILLIC DASIA PNEUMATA
U+0486	230	COMBINING CYRILLIC PSILI PNEUMATA
U+0592	230	HEBREW ACCENT SEGOL
U+0593	230	HEBREW ACCENT SHALSHELET
U+0594	230	HEBREW ACCENT ZAQEF QATAN
U+0595	230	HEBREW ACCENT ZAQEF GADOL
U+0597	230	HEBREW ACCENT REVIA
U+0598	230	HEBREW ACCENT ZARQA
U+0599	230	HEBREW ACCENT PASHTA
U+059C	230	HEBREW ACCENT GERESH
U+059D	230	HEBREW ACCENT GERESH MUQDAM
U+059E	230	HEBREW ACCENT GERSHAYIM
U+059F	230	HEBREW ACCENT QARNEY PARA
U+05A0	230	HEBREW ACCENT TELISHA GEDOLA
U+05A1	230	HEBREW ACCENT PAZER
U+05A8	230	HEBREW ACCENT QADMA
U+05A9	230	HEBREW ACCENT TELISHA QETANA
U+05AB	230	HEBREW ACCENT OLE

Table 4-3. Combining Classes (Continued)

Code	Class	Name
U+05AC	230	HEBREW ACCENT ILUY
U+05AE	230	HEBREW ACCENT ZINOR
U+05AF	230	HEBREW MARK MASORA CIRCLE
U+05C4	230	HEBREW MARK UPPER DOT
U+06D7	230	ARABIC SMALL HIGH LIGATURE QAF WITH LAM WITH ALEF MAKSURA
U+06D8	230	ARABIC SMALL HIGH MEEM INITIAL FORM
U+06D9	230	ARABIC SMALL HIGH LAM ALEF
U+06DA	230	ARABIC SMALL HIGH JEEM
U+06DB	230	ARABIC SMALL HIGH THREE DOTS
U+06DC	230	ARABIC SMALL HIGH SEEN
U+06DF	230	ARABIC SMALL HIGH ROUNDED ZERO
U+06E0	230	ARABIC SMALL HIGH UPRIGHT RECTANGULAR ZERO
U+06E1	230	ARABIC SMALL HIGH DOTLESS HEAD OF KHAH
U+06E2	230	ARABIC SMALL HIGH MEEM ISOLATED FORM
U+06E4	230	ARABIC SMALL HIGH MADDA
U+06E7	230	ARABIC SMALL HIGH YEH
U+06E8	230	ARABIC SMALL HIGH NOON
U+06EB	230	ARABIC EMPTY CENTRE HIGH STOP
U+06EC	230	ARABIC ROUNDED HIGH STOP WITH FILLED CENTRE
U+0953	230	DEVANAGARI GRAVE ACCENT
U+0954	230	DEVANAGARI ACUTE ACCENT
U+0B56	230	ORIYA AI LENGTH MARK
U+0F82	230	TIBETAN SIGN NYI ZLA NAA DA
U+0F83	230	TIBETAN SIGN SNA LDAN
U+0F86	230	TIBETAN MARK LCI RTAGS
U+0F87	230	TIBETAN MARK YANG RTAGS
U+0F88	230	TIBETAN MARK LCE TSA CAN
U+0F89	230	TIBETAN MARK MCHU CAN
U+0F8A	230	TIBETAN MARK GRU CAN RGYINGS
U+0F8B	230	TIBETAN MARK GRU MED RGYINGS
U+20D0	230	COMBINING LEFT HARPOON ABOVE
U+20D1	230	COMBINING RIGHT HARPOON ABOVE
U+20D4	230	COMBINING ANTICLOCKWISE ARROW ABOVE
U+20D5	230	COMBINING CLOCKWISE ARROW ABOVE
U+20D6	230	COMBINING LEFT ARROW ABOVE
U+20D7	230	COMBINING RIGHT ARROW ABOVE
U+20DB	230	COMBINING THREE DOTS ABOVE
U+20DC	230	COMBINING FOUR DOTS ABOVE
U+FE20	230	COMBINING LIGATURE LEFT HALF
U+FE21	230	COMBINING LIGATURE RIGHT HALF
U+FE22	230	COMBINING DOUBLE TILDE LEFT HALF
U+FE23	230	COMBINING DOUBLE TILDE RIGHT HALF
<i>Above Right</i>		
U+0315	232	COMBINING COMMA ABOVE RIGHT
U+031A	232	COMBINING LEFT ANGLE ABOVE
U+302C	232	IDEOGRAPHIC DEPARTING TONE MARK
<i>Double Above</i>		
U+0360	234	COMBINING DOUBLE TILDE
U+0361	234	COMBINING DOUBLE INVERTED BREVE

4.3 Directionality

Directional behavior is interpreted according to the Unicode Bidirectional Algorithm (see Section 3.11, *Bidirectional Behavior*). For this purpose, all characters of the Unicode Standard possess a normative *directional* type. The directional types left-to-right and right-to-left are called *strong types*, and characters of these types are called strong directional characters. Left-to-right types include most alphabetic and syllabic characters, as well as all Han

ideographic characters. Right-to-left types include Arabic, Hebrew, and punctuation specific to those scripts. In addition, the Unicode Bidirectional Algorithm also uses *weak types* and *neutrals*.

As with other character properties, Compatibility Area characters have the same directional properties as the characters they map to. The directional type of all unassigned code values is not defined. This is also true of unassigned code values falling within the ranges used in the following table; for brevity, not all unassigned code values in the ranges are called out separately where there are gaps. As characters are assigned in future versions of the Unicode Standard, their properties will be documented.

It is recommended that unassigned code values be given the property of neutral for compatibility with future versions of this standard where these code points are assigned. That choice is most likely to produce acceptable rendering.

Table 4-4 gives the directional type of Unicode characters.

Table 4-4. Bidirectional Character Types

Type	Sym	Category	Description	Values
Strong	L	<i>Left-to-Right</i>	Miscellaneous	U+0026, U+0040
			Latin letters	U+0041→U+005A, U+0061→U+007A, U+00C0→U+00D6, U+00D8→U+00F6, U+00F8→U+00FF
			Extended Latin→Modifier Letters	U+0100→U+02FF
			Combining Diacriticals	U+0300→U+036F
			Greek→Armenian	U+0370→U+058F
			Devanagari→Thai Character Phinthu	U+0900→U+0E3A
			Thai Character Sara E→Hangul Jamo	U+0E40→U+11FF
			Latin Extended Additional→Greek Extended	U+1E00→U+1FFF
			Left-to-Right Mark	U+200E
			Symbol Diacritics	U+20D0→U+20FF
			Roman Numerals	U+2160→U+2182
			Hiragana→Han	U+3040→U+9FFF
	R	<i>Right-to-Left</i>	Arabic and Hebrew	U+0590→U+065F, U+066D→U+06EF
			Right-to-Left Mark	U+200F
Weak	EN	<i>European Number</i>	European Digits	U+0030→U+0039
			Eastern Arabic-Indic Digits	U+06F0→U+06F9
			Superscript and Subscript Digits	U+00B2→U+00B3, U+00B9, U+2070, U+2074→U+2079, U+2080→U+2089
	ES	<i>European Number Separator</i>	Full Stop (Period)	U+002E
			Solidus (Slash)	U+002F
			Figure Space	U+2007
	ET	<i>European Number Terminator</i>	Number Sign	U+0023

Table 4-4. Bidirectional Character Types (Continued)

Type	Sym	Category	Description	Values
			Currency symbols	U+0024, U+00A2→U+00A5, U+0E3F, U+20A0→U+20CF
			Percents	U+0025, U+066A
			Plus Sign	U+002B
			Hyphen-Minus	U+002D
			Degree	U+00B0
			Plus-Minus	U+00B1
			Prime (Minute)	U+2032
			Double Prime (Second)	U+2033
			Superscript Plus and Minus	U+207A, U+207B
			Subscript Plus and Minus	U+208A, U+208B
			Minus Sign	U+2212
			Minus-Plus	U+2213
	AN	Arabic Number	Arabic-Indic Digits	U+0660→U+0669
			Arabic Decimal and Thousands Separators	U+066B, U+066C
	CS	Common Number Separator	Comma	U+002C
			Colon	U+003A
Separators	B	Block Separator	Line Separator	U+2028
			Paragraph Separator	U+2029
	S	Segment Separator	Horizontal Tab	U+0009
Neutrals	WS	Whitespace	Space	U+0020
			No-Break Space	U+00A0
			General Punctuation Spaces	U+2000→U+2006, U+2008→U+200B, U+3000, U+FEFF
	ON	Other Neutrals	Punctuation, Symbols	All other characters

The definition of control code semantics is generally outside of the scope of the Unicode Standard (see Section 2.6, *Controls and Control Sequences*). Implementers should interpret the type of characters such as *carriage return*, *line feed*, *group separator*, and so on according to the closest semantics to the types given here, such as interpreting *carriage return* (when used as paragraph separator) as being a block separator. However, U+0009 HORIZONTAL TAB is interpreted as a segment separator, which indicates that in a line containing tabs, the tab-delimited segments go in the base level direction (see Section 3.11, *Bidirectional Behavior*).

Since horizontal Han ideographic characters are generally arranged from left to right when written in horizontally oriented lines, they have the left-to-right directional character type. (When they need to be written from right to left in horizontal lines, their direction can be overridden, as discussed in the Section 3.11, *Bidirectional Behavior*.)

4.4 Jamo Short Names

The *jamo short name* is the normative property of the Unicode conjoining Hangul *jamo* characters. These short names are used to determine the character names that are derived when decomposing Hangul syllables into their decomposition sequence.

Table 4-5. Jamo Short Names

Value	Short Name	Unicode Name
U+1100	G	HANGUL CHOSEONG KIYEOK
U+1101	GG	HANGUL CHOSEONG SSANGKIYEOK
U+1102	N	HANGUL CHOSEONG NIEUN
U+1103	D	HANGUL CHOSEONG TIKEUT
U+1104	DD	HANGUL CHOSEONG SSANGTIKEUT
U+1105	L	HANGUL CHOSEONG RIEUL
U+1106	M	HANGUL CHOSEONG MIEUM
U+1107	B	HANGUL CHOSEONG PIEUP
U+1108	BB	HANGUL CHOSEONG SSANGPIEUP
U+1109	S	HANGUL CHOSEONG SIOS
U+110A	SS	HANGUL CHOSEONG SSANGSIOS
U+110B		HANGUL CHOSEONG IEUNG
U+110C	J	HANGUL CHOSEONG CIEUC
U+110D	JJ	HANGUL CHOSEONG SSANGCIEUC
U+110E	C	HANGUL CHOSEONG CHIEUCH
U+110F	K	HANGUL CHOSEONG KHIEUKH
U+1110	T	HANGUL CHOSEONG THIEUTH
U+1111	P	HANGUL CHOSEONG PHIEUPH
U+1112	H	HANGUL CHOSEONG HIEUH
U+1161	A	HANGUL JUNGSEONG A
U+1162	AE	HANGUL JUNGSEONG AE
U+1163	YA	HANGUL JUNGSEONG YA
U+1164	YAE	HANGUL JUNGSEONG YAE
U+1165	EO	HANGUL JUNGSEONG EO
U+1166	E	HANGUL JUNGSEONG E
U+1167	YEO	HANGUL JUNGSEONG YEO
U+1168	YE	HANGUL JUNGSEONG YE
U+1169	O	HANGUL JUNGSEONG O
U+116A	WA	HANGUL JUNGSEONG WA
U+116B	WAE	HANGUL JUNGSEONG WAE
U+116C	OE	HANGUL JUNGSEONG OE
U+116D	YO	HANGUL JUNGSEONG YO
U+116E	U	HANGUL JUNGSEONG U
U+116F	WEO	HANGUL JUNGSEONG WEO
U+1170	WE	HANGUL JUNGSEONG WE
U+1171	WI	HANGUL JUNGSEONG WI
U+1172	YU	HANGUL JUNGSEONG YU
U+1173	EU	HANGUL JUNGSEONG EU
U+1174	YI	HANGUL JUNGSEONG YI
U+1175	I	HANGUL JUNGSEONG I
U+11A8	G	HANGUL JONGSEONG KIYEOK
U+11A9	GG	HANGUL JONGSEONG SSANGKIYEOK
U+11AA	GS	HANGUL JONGSEONG KIYEOK-SIOS
U+11AB	N	HANGUL JONGSEONG NIEUN
U+11AC	NJ	HANGUL JONGSEONG NIEUN-CIEUC
U+11AD	NH	HANGUL JONGSEONG NIEUN-HIEUH
U+11AE	D	HANGUL JONGSEONG TIKEUT
U+11AF	L	HANGUL JONGSEONG RIEUL
U+11B0	LG	HANGUL JONGSEONG RIEUL-KIYEOK

Table 4-5. Jamo Short Names (Continued)

Value	Short Name	Unicode Name
U+11B1	LM	HANGUL JONGSEONG RIEUL-MIEUM
U+11B2	LB	HANGUL JONGSEONG RIEUL-PIEUP
U+11B3	LS	HANGUL JONGSEONG RIEUL-SIOS
U+11B4	LT	HANGUL JONGSEONG RIEUL-THIEUTH
U+11B5	LP	HANGUL JONGSEONG RIEUL-PHIEUPH
U+11B6	LH	HANGUL JONGSEONG RIEUL-HIEUH
U+11B7	M	HANGUL JONGSEONG MIEUM
U+11B8	B	HANGUL JONGSEONG PIEUP
U+11B9	BS	HANGUL JONGSEONG PIEUP-SIOS
U+11BA	S	HANGUL JONGSEONG SIOS
U+11BB	SS	HANGUL JONGSEONG SSANGSIOS
U+11BC	NG	HANGUL JONGSEONG IEUNG
U+11BD	J	HANGUL JONGSEONG CIEUC
U+11BE	C	HANGUL JONGSEONG CHIEUCH
U+11BF	K	HANGUL JONGSEONG KHIEUKH
U+11C0	T	HANGUL JONGSEONG THIEUTH
U+11C1	P	HANGUL JONGSEONG PHIEUPH
U+11C2	H	HANGUL JONGSEONG HIEUH

4.5 Letters

Letter. This is an informative property of characters that are used to write words. This includes characters such as capital letters, small letters, ideographs, hangul, and spacing modifier letters. Combining marks generally assume the letter property of the preceding base character. For example, when searching for word boundaries, combining characters don't break from previous letters. The letter property mappings can be obtained from the CD-ROM accompanying the standard.

There are two informative character properties that specify subsets of all letters, alphabetic and ideographic.

Alphabetic. This is the property of the primary units of alphabets and syllabaries, both combining and non-combining. It also includes modifier letters and letterlike symbols that are compatibility equivalents to alphabetic letters. The following is a list of all characters in the Unicode Standard which have the alphabetic property (see *Section 5.14, Identifiers*).

```

0041..005A, 0061..007A, 00AA      , 00B5      , 00BA      , 00C0..00D6,
00D8..00F6, 00F8..01F5, 01FA..0217, 0250..02A8, 02B0..02B8, 02BB..02C1,
02E0..02E4, 037A      , 0386      , 0388..038A, 038C      , 038E..03A1,
03A3..03CE, 03D0..03D6, 03DA      , 03DC      , 03DE      , 03E0      ,
03E2..03F3, 0401..040C, 040E..044F, 0451..045C, 045E..0481, 0490..04C4,
04C7..04C8, 04CB..04CC, 04D0..04EB, 04EE..04F5, 04F8..04F9, 0531..0556,
0559..055A, 0561..0587, 05D0..05EA, 05F0..05F2, 0621..063A, 0641..0652,
0670..06B7, 06BA..06BE, 06C0..06CE, 06D0..06D3, 06D5..06DC, 06E1..06E8,
06ED      , 0901..0903, 0905..0939, 093D..094C, 0958..0963, 0981..0983,
0985..098C, 098F..0990, 0993..09A8, 09AA..09B0, 09B2      , 09B6..09B9,
09BE..09C4, 09C7..09C8, 09CB..09CC, 09D7      , 09DC..09DD, 09DF..09E3,
09F0..09F1, 0A02      , 0A05..0A0A, 0A0F..0A10, 0A13..0A28, 0A2A..0A30,
0A32..0A33, 0A35..0A36, 0A38..0A39, 0A3E..0A42, 0A47..0A48, 0A4B..0A4C,
0A8F..0A91, 0A93..0AA8, 0AAA..0AB0, 0AB2..0AB3, 0AB5..0AB9, 0ABD..0AC5,
0AC7..0AC9, 0ACB..0ACC, 0AE0      , 0B01..0B03, 0B05..0B0C, 0B0F..0B10,
0B13..0B28, 0B2A..0B30, 0B32..0B33, 0B36..0B39, 0B3D..0B43, 0B47..0B48,
0B4B..0B4C, 0B56..0B57, 0B5C..0B5D, 0B5F..0B61, 0B82..0B83, 0B85..0B8A,
0B8E..0B90, 0B92..0B95, 0B99..0B9A, 0B9C      , 0B9E..0B9F, 0BA3..0BA4,

```

```

0BA8..0BAA, 0BAE..0BB5, 0BB7..0BB9, 0BBE..0BC2, 0BC6..0BC8, 0BCA..0BCC,
0BD7, 0C01..0C03, 0C05..0C0C, 0C0E..0C10, 0C12..0C28, 0C2A..0C33,
0C35..0C39, 0C3E..0C44, 0C46..0C48, 0C4A..0C4C, 0C55..0C56, 0C60..0C61,
0C82..0C83, 0C85..0C8C, 0C8E..0C90, 0C92..0CA8, 0CAA..0CB3, 0CB5..0CB9,
0CBE..0CC4, 0CC6..0CC8, 0CCA..0CCC, 0CD5..0CD6, 0CDE, 0CE0..0CE1,
0D02..0D03, 0D05..0D0C, 0D0E..0D10, 0D12..0D28, 0D2A..0D39, 0D3E..0D43,
0D46..0D48, 0D4A..0D4C, 0D57, 0D60..0D61, 0E01..0E2E, 0E30..0E3A,
0E40..0E45, 0E47, 0E4D, 0E81..0E82, 0E84, 0E87..0E88,
0E8A, 0E8D, 0E94..0E97, 0E99..0E9F, 0EA1..0EA3, 0EA5,
0EA7, 0EAA..0EAB, 0EAD..0EAE, 0EB0..0EB9, 0EBB..0EBD, 0EC0..0EC4,
0ECD, 0EDC..0EDD, 0F40..0F47, 0F49..0F69, 0F71..0F81, 0F90..0F95,
0F97, 0F99..0FAD, 0FB1..0FB7, 0FB9, 10A0..10C5, 10D0..10F6,
1100..1159, 115F..11A2, 11A8..11F9, 1E00..1E9B, 1EA0..1EF9, 1F00..1F15,
1F18..1F1D, 1F20..1F45, 1F48..1F4D, 1F50..1F57, 1F59, 1F5B,
1F5D, 1F5F..1F7D, 1F80..1FB4, 1FB6..1FBC, 1FBE, 1FC2..1FC4,
1FC6..1FCC, 1FD0..1FD3, 1FD6..1FDB, 1FE0..1FEC, 1FF2..1FF4, 1FF6..1FFC,
207F, 2102, 2107, 210A..2113, 2115, 2118..211D,
2124, 2126, 2128, 212A..212D, 212F..2131, 2133..2138,
2160..2182, 3041..3094, 30A1..30FA, 3105..312C, 3131..318E, AC00..D7A3,
FB00..FB06, FB13..FB17, FB1F..FB28, FB2A..FB36, FB38..FB3C, FB3E,
FB40..FB41, FB43..FB44, FB46..FBB1, FBD3..FD3D, FD50..FD8F, FD92..FDC7,
FDF0..FDFB, FE70..FE72, FE74, FE76..FEFC, FF21..FF3A, FF41..FF5A,
FF66..FF6F, FF71..FF9D, FFA0..FFBE, FFC2..FFC7, FFCA..FFCF, FFD2..FFD7,
FFDA..FFDC.

```

Ideographic. This is the property of the Unified Han Set and other Han characters. The following is a list of all characters in the Unicode Standard which have the ideographic property. See Section 5.14, *Identifiers*.

```
4E00..9FA5, F900..FA2D, 3007, 3021..3029
```

4.6 Numeric Value

Numeric value is a normative property of characters that represent *numbers*. This includes characters such as fractions, subscripts, superscripts, Roman numerals, currency numerators, encircled numbers, and script-specific digits. In many traditional numbering systems, letters are used with a numeric value. Examples include Greek and Hebrew letters, and Latin letters used in outlines (II.A.1.b). These are special cases and are not included here as numbers.

Decimal digits form a large subcategory of numbers consisting of those digits that can be used to form decimal-radix numbers. This includes script-specific digits. Decimal digits do not include characters such as Roman numerals ($1 + 5 = 15 = \text{fifteen}$, but $I + V = IV = \text{four}$), subscripts, or superscripts. Numbers other than decimal digits can be used in numerical expressions, but it is up to the user to determine the specialized uses.

The Unicode Standard assigns distinct codes to the forms of digits that are specific to a given script or language. Examples are the digits used with the Arabic script, Chinese numbers, or those of the Indic languages. For naming conventions, see the introduction to the Arabic subsection of Section 6.1, *General Scripts Area*.

Table 4-6 gives the numeric values of Unicode characters that can represent numbers. (Some CJK ideographs also have numeric values, but are not included in Table 4-6 in this version of the Unicode Standard.)

Table 4-6. Numeric Properties

Value	Decimal	Number	Name
U+0030	✓	0	DIGIT ZERO
U+0031	✓	1	DIGIT ONE
U+0032	✓	2	DIGIT TWO
U+0033	✓	3	DIGIT THREE

Table 4-6. Numeric Properties (Continued)

Value	Decimal	Number	Name
U+0034	✓	4	DIGIT FOUR
U+0035	✓	5	DIGIT FIVE
U+0036	✓	6	DIGIT SIX
U+0037	✓	7	DIGIT SEVEN
U+0038	✓	8	DIGIT EIGHT
U+0039	✓	9	DIGIT NINE
U+00B2		2	SUPERSCRIPT TWO
U+00B3		3	SUPERSCRIPT THREE
U+00B9		1	SUPERSCRIPT ONE
U+00BC		1/4	VULGAR FRACTION ONE QUARTER
U+00BD		1/2	VULGAR FRACTION ONE HALF
U+00BE		3/4	VULGAR FRACTION THREE QUARTERS
U+0660	✓	0	ARABIC-INDIC DIGIT ZERO
U+0661	✓	1	ARABIC-INDIC DIGIT ONE
U+0662	✓	2	ARABIC-INDIC DIGIT TWO
U+0663	✓	3	ARABIC-INDIC DIGIT THREE
U+0664	✓	4	ARABIC-INDIC DIGIT FOUR
U+0665	✓	5	ARABIC-INDIC DIGIT FIVE
U+0666	✓	6	ARABIC-INDIC DIGIT SIX
U+0667	✓	7	ARABIC-INDIC DIGIT SEVEN
U+0668	✓	8	ARABIC-INDIC DIGIT EIGHT
U+0669	✓	9	ARABIC-INDIC DIGIT NINE
U+06F0	✓	0	EXTENDED ARABIC-INDIC DIGIT ZERO
U+06F1	✓	1	EXTENDED ARABIC-INDIC DIGIT ONE
U+06F2	✓	2	EXTENDED ARABIC-INDIC DIGIT TWO
U+06F3	✓	3	EXTENDED ARABIC-INDIC DIGIT THREE
U+06F4	✓	4	EXTENDED ARABIC-INDIC DIGIT FOUR
U+06F5	✓	5	EXTENDED ARABIC-INDIC DIGIT FIVE
U+06F6	✓	6	EXTENDED ARABIC-INDIC DIGIT SIX
U+06F7	✓	7	EXTENDED ARABIC-INDIC DIGIT SEVEN
U+06F8	✓	8	EXTENDED ARABIC-INDIC DIGIT EIGHT
U+06F9	✓	9	EXTENDED ARABIC-INDIC DIGIT NINE
U+0966	✓	0	DEVANAGARI DIGIT ZERO
U+0967	✓	1	DEVANAGARI DIGIT ONE
U+0968	✓	2	DEVANAGARI DIGIT TWO
U+0969	✓	3	DEVANAGARI DIGIT THREE
U+096A	✓	4	DEVANAGARI DIGIT FOUR
U+096B	✓	5	DEVANAGARI DIGIT FIVE
U+096C	✓	6	DEVANAGARI DIGIT SIX
U+096D	✓	7	DEVANAGARI DIGIT SEVEN
U+096E	✓	8	DEVANAGARI DIGIT EIGHT
U+096F	✓	9	DEVANAGARI DIGIT NINE
U+09E6	✓	0	BENGALI DIGIT ZERO
U+09E7	✓	1	BENGALI DIGIT ONE
U+09E8	✓	2	BENGALI DIGIT TWO
U+09E9	✓	3	BENGALI DIGIT THREE
U+09EA	✓	4	BENGALI DIGIT FOUR
U+09EB	✓	5	BENGALI DIGIT FIVE
U+09EC	✓	6	BENGALI DIGIT SIX
U+09ED	✓	7	BENGALI DIGIT SEVEN
U+09EE	✓	8	BENGALI DIGIT EIGHT
U+09EF	✓	9	BENGALI DIGIT NINE
U+09F4		1	BENGALI CURRENCY NUMERATOR ONE
U+09F5		2	BENGALI CURRENCY NUMERATOR TWO
U+09F6		3	BENGALI CURRENCY NUMERATOR THREE
U+09F7		4	BENGALI CURRENCY NUMERATOR FOUR
U+09F8		-	BENGALI CURRENCY NUMERATOR ONE LESS THAN THE DENOMINATOR
U+09F9		16	BENGALI CURRENCY DENOMINATOR SIXTEEN
U+0A66	✓	0	GURMUKHI DIGIT ZERO
U+0A67	✓	1	GURMUKHI DIGIT ONE

Table 4-6. Numeric Properties (Continued)

Value	Decimal	Number	Name
U+0A68	✓	2	GURMUKHI DIGIT TWO
U+0A69	✓	3	GURMUKHI DIGIT THREE
U+0A6A	✓	4	GURMUKHI DIGIT FOUR
U+0A6B	✓	5	GURMUKHI DIGIT FIVE
U+0A6C	✓	6	GURMUKHI DIGIT SIX
U+0A6D	✓	7	GURMUKHI DIGIT SEVEN
U+0A6E	✓	8	GURMUKHI DIGIT EIGHT
U+0A6F	✓	9	GURMUKHI DIGIT NINE
U+0AE6	✓	0	GUJARATI DIGIT ZERO
U+0AE7	✓	1	GUJARATI DIGIT ONE
U+0AE8	✓	2	GUJARATI DIGIT TWO
U+0AE9	✓	3	GUJARATI DIGIT THREE
U+0AEA	✓	4	GUJARATI DIGIT FOUR
U+0AEB	✓	5	GUJARATI DIGIT FIVE
U+0AEC	✓	6	GUJARATI DIGIT SIX
U+0AED	✓	7	GUJARATI DIGIT SEVEN
U+0AEE	✓	8	GUJARATI DIGIT EIGHT
U+0AEF	✓	9	GUJARATI DIGIT NINE
U+0B66	✓	0	ORIYA DIGIT ZERO
U+0B67	✓	1	ORIYA DIGIT ONE
U+0B68	✓	2	ORIYA DIGIT TWO
U+0B69	✓	3	ORIYA DIGIT THREE
U+0B6A	✓	4	ORIYA DIGIT FOUR
U+0B6B	✓	5	ORIYA DIGIT FIVE
U+0B6C	✓	6	ORIYA DIGIT SIX
U+0B6D	✓	7	ORIYA DIGIT SEVEN
U+0B6E	✓	8	ORIYA DIGIT EIGHT
U+0B6F	✓	9	ORIYA DIGIT NINE
U+0BE7	✓	1	TAMIL DIGIT ONE
U+0BE8	✓	2	TAMIL DIGIT TWO
U+0BE9	✓	3	TAMIL DIGIT THREE
U+0BEA	✓	4	TAMIL DIGIT FOUR
U+0BEB	✓	5	TAMIL DIGIT FIVE
U+0BEC	✓	6	TAMIL DIGIT SIX
U+0BED	✓	7	TAMIL DIGIT SEVEN
U+0BEE	✓	8	TAMIL DIGIT EIGHT
U+0BEF	✓	9	TAMIL DIGIT NINE
U+0BF0		10	TAMIL NUMBER TEN
U+0BF1		100	TAMIL NUMBER ONE HUNDRED
U+0BF2		1000	TAMIL NUMBER ONE THOUSAND
U+0C66	✓	0	TELUGU DIGIT ZERO
U+0C67	✓	1	TELUGU DIGIT ONE
U+0C68	✓	2	TELUGU DIGIT TWO
U+0C69	✓	3	TELUGU DIGIT THREE
U+0C6A	✓	4	TELUGU DIGIT FOUR
U+0C6B	✓	5	TELUGU DIGIT FIVE
U+0C6C	✓	6	TELUGU DIGIT SIX
U+0C6D	✓	7	TELUGU DIGIT SEVEN
U+0C6E	✓	8	TELUGU DIGIT EIGHT
U+0C6F	✓	9	TELUGU DIGIT NINE
U+0CE6	✓	0	KANNADA DIGIT ZERO
U+0CE7	✓	1	KANNADA DIGIT ONE
U+0CE8	✓	2	KANNADA DIGIT TWO
U+0CE9	✓	3	KANNADA DIGIT THREE
U+0CEA	✓	4	KANNADA DIGIT FOUR
U+0CEB	✓	5	KANNADA DIGIT FIVE
U+0CEC	✓	6	KANNADA DIGIT SIX
U+0CED	✓	7	KANNADA DIGIT SEVEN
U+0CEE	✓	8	KANNADA DIGIT EIGHT
U+0CEF	✓	9	KANNADA DIGIT NINE
U+0D66	✓	0	MALAYALAM DIGIT ZERO

Table 4-6. Numeric Properties (Continued)

Value	Decimal	Number	Name
U+0D67	✓	1	MALAYALAM DIGIT ONE
U+0D68	✓	2	MALAYALAM DIGIT TWO
U+0D69	✓	3	MALAYALAM DIGIT THREE
U+0D6A	✓	4	MALAYALAM DIGIT FOUR
U+0D6B	✓	5	MALAYALAM DIGIT FIVE
U+0D6C	✓	6	MALAYALAM DIGIT SIX
U+0D6D	✓	7	MALAYALAM DIGIT SEVEN
U+0D6E	✓	8	MALAYALAM DIGIT EIGHT
U+0D6F	✓	9	MALAYALAM DIGIT NINE
U+0E50	✓	0	THAI DIGIT ZERO
U+0E51	✓	1	THAI DIGIT ONE
U+0E52	✓	2	THAI DIGIT TWO
U+0E53	✓	3	THAI DIGIT THREE
U+0E54	✓	4	THAI DIGIT FOUR
U+0E55	✓	5	THAI DIGIT FIVE
U+0E56	✓	6	THAI DIGIT SIX
U+0E57	✓	7	THAI DIGIT SEVEN
U+0E58	✓	8	THAI DIGIT EIGHT
U+0E59	✓	9	THAI DIGIT NINE
U+0ED0	✓	0	LAO DIGIT ZERO
U+0ED1	✓	1	LAO DIGIT ONE
U+0ED2	✓	2	LAO DIGIT TWO
U+0ED3	✓	3	LAO DIGIT THREE
U+0ED4	✓	4	LAO DIGIT FOUR
U+0ED5	✓	5	LAO DIGIT FIVE
U+0ED6	✓	6	LAO DIGIT SIX
U+0ED7	✓	7	LAO DIGIT SEVEN
U+0ED8	✓	8	LAO DIGIT EIGHT
U+0ED9	✓	9	LAO DIGIT NINE
U+0F20	✓	0	TIBETAN DIGIT ZERO
U+0F21	✓	1	TIBETAN DIGIT ONE
U+0F22	✓	2	TIBETAN DIGIT TWO
U+0F23	✓	3	TIBETAN DIGIT THREE
U+0F2D	✓	4	TIBETAN DIGIT FOUR
U+0F25	✓	5	TIBETAN DIGIT FIVE
U+0F26	✓	6	TIBETAN DIGIT SIX
U+0F27	✓	7	TIBETAN DIGIT SEVEN
U+0F28	✓	8	TIBETAN DIGIT EIGHT
U+0F29	✓	9	TIBETAN DIGIT NINE
U+0F2A		1/2	TIBETAN DIGIT HALF ONE
U+0F2B		3/2	TIBETAN DIGIT HALF TWO
U+0F2C		5/2	TIBETAN DIGIT HALF THREE
U+0F2D		7/2	TIBETAN DIGIT HALF FOUR
U+0F2E		9/2	TIBETAN DIGIT HALF FIVE
U+0F2F		11/2	TIBETAN DIGIT HALF SIX
U+0F30		13/2	TIBETAN DIGIT HALF SEVEN
U+0F31		15/2	TIBETAN DIGIT HALF EIGHT
U+0F32		17/2	TIBETAN DIGIT HALF NINE
U+0F33		-1/2	TIBETAN DIGIT HALF ZERO
U+2070		0	SUPERSCRIPIT ZERO
U+2074		4	SUPERSCRIPIT FOUR
U+2075		5	SUPERSCRIPIT FIVE
U+2076		6	SUPERSCRIPIT SIX
U+2077		7	SUPERSCRIPIT SEVEN
U+2078		8	SUPERSCRIPIT EIGHT
U+2079		9	SUPERSCRIPIT NINE
U+2080		0	SUBSCRIPT ZERO
U+2081		1	SUBSCRIPT ONE
U+2082		2	SUBSCRIPT TWO
U+2083		3	SUBSCRIPT THREE
U+2084		4	SUBSCRIPT FOUR

Table 4-6. Numeric Properties (Continued)

Value	Decimal	Number	Name
U+2085		5	SUBSCRIPT FIVE
U+2086		6	SUBSCRIPT SIX
U+2087		7	SUBSCRIPT SEVEN
U+2088		8	SUBSCRIPT EIGHT
U+2089		9	SUBSCRIPT NINE
U+2153		1/3	VULGAR FRACTION ONE THIRD
U+2154		2/3	VULGAR FRACTION TWO THIRDS
U+2155		1/5	VULGAR FRACTION ONE FIFTH
U+2156		2/5	VULGAR FRACTION TWO FIFTHS
U+2157		3/5	VULGAR FRACTION THREE FIFTHS
U+2158		4/5	VULGAR FRACTION FOUR FIFTHS
U+2159		1/6	VULGAR FRACTION ONE SIXTH
U+215A		5/6	VULGAR FRACTION FIVE SIXTHS
U+215B		1/8	VULGAR FRACTION ONE EIGHTH
U+215C		3/8	VULGAR FRACTION THREE EIGHTHS
U+215D		5/8	VULGAR FRACTION FIVE EIGHTHS
U+215E		7/8	VULGAR FRACTION SEVEN EIGHTHS
U+215F		1	FRACTION NUMERATOR ONE
U+2160		1	ROMAN NUMERAL ONE
U+2161		2	ROMAN NUMERAL TWO
U+2162		3	ROMAN NUMERAL THREE
U+2163		4	ROMAN NUMERAL FOUR
U+2164		5	ROMAN NUMERAL FIVE
U+2165		6	ROMAN NUMERAL SIX
U+2166		7	ROMAN NUMERAL SEVEN
U+2167		8	ROMAN NUMERAL EIGHT
U+2168		9	ROMAN NUMERAL NINE
U+2169		10	ROMAN NUMERAL TEN
U+216A		11	ROMAN NUMERAL ELEVEN
U+216B		12	ROMAN NUMERAL TWELVE
U+216C		50	ROMAN NUMERAL FIFTY
U+216D		100	ROMAN NUMERAL ONE HUNDRED
U+216E		500	ROMAN NUMERAL FIVE HUNDRED
U+216F		1000	ROMAN NUMERAL ONE THOUSAND
U+2170		1	SMALL ROMAN NUMERAL ONE
U+2171		2	SMALL ROMAN NUMERAL TWO
U+2172		3	SMALL ROMAN NUMERAL THREE
U+2173		4	SMALL ROMAN NUMERAL FOUR
U+2174		5	SMALL ROMAN NUMERAL FIVE
U+2175		6	SMALL ROMAN NUMERAL SIX
U+2176		7	SMALL ROMAN NUMERAL SEVEN
U+2177		8	SMALL ROMAN NUMERAL EIGHT
U+2178		9	SMALL ROMAN NUMERAL NINE
U+2179		10	SMALL ROMAN NUMERAL TEN
U+217A		11	SMALL ROMAN NUMERAL ELEVEN
U+217B		12	SMALL ROMAN NUMERAL TWELVE
U+217C		50	SMALL ROMAN NUMERAL FIFTY
U+217D		100	SMALL ROMAN NUMERAL ONE HUNDRED
U+217E		500	SMALL ROMAN NUMERAL FIVE HUNDRED
U+217F		1000	SMALL ROMAN NUMERAL ONE THOUSAND
U+2180		1000	ROMAN NUMERAL ONE THOUSAND C D
U+2181		5000	ROMAN NUMERAL FIVE THOUSAND
U+2182		10000	ROMAN NUMERAL TEN THOUSAND
U+2460		1	CIRCLED DIGIT ONE
U+2461		2	CIRCLED DIGIT TWO
U+2462		3	CIRCLED DIGIT THREE
U+2463		4	CIRCLED DIGIT FOUR
U+2464		5	CIRCLED DIGIT FIVE
U+2465		6	CIRCLED DIGIT SIX
U+2466		7	CIRCLED DIGIT SEVEN
U+2467		8	CIRCLED DIGIT EIGHT

Table 4-6. Numeric Properties (Continued)

Value	Decimal	Number	Name
U+2468		9	CIRCLED DIGIT NINE
U+2469		10	CIRCLED NUMBER TEN
U+246A		11	CIRCLED NUMBER ELEVEN
U+246B		12	CIRCLED NUMBER TWELVE
U+246C		13	CIRCLED NUMBER THIRTEEN
U+246D		14	CIRCLED NUMBER FOURTEEN
U+246E		15	CIRCLED NUMBER FIFTEEN
U+246F		16	CIRCLED NUMBER SIXTEEN
U+2470		17	CIRCLED NUMBER SEVENTEEN
U+2471		18	CIRCLED NUMBER EIGHTEEN
U+2472		19	CIRCLED NUMBER NINETEEN
U+2473		20	CIRCLED NUMBER TWENTY
U+2474		1	PARENTHESIZED DIGIT ONE
U+2475		2	PARENTHESIZED DIGIT TWO
U+2476		3	PARENTHESIZED DIGIT THREE
U+2477		4	PARENTHESIZED DIGIT FOUR
U+2478		5	PARENTHESIZED DIGIT FIVE
U+2479		6	PARENTHESIZED DIGIT SIX
U+247A		7	PARENTHESIZED DIGIT SEVEN
U+247B		8	PARENTHESIZED DIGIT EIGHT
U+247C		9	PARENTHESIZED DIGIT NINE
U+247D		10	PARENTHESIZED NUMBER TEN
U+247E		11	PARENTHESIZED NUMBER ELEVEN
U+247F		12	PARENTHESIZED NUMBER TWELVE
U+2480		13	PARENTHESIZED NUMBER THIRTEEN
U+2481		14	PARENTHESIZED NUMBER FOURTEEN
U+2482		15	PARENTHESIZED NUMBER FIFTEEN
U+2483		16	PARENTHESIZED NUMBER SIXTEEN
U+2484		17	PARENTHESIZED NUMBER SEVENTEEN
U+2485		18	PARENTHESIZED NUMBER EIGHTEEN
U+2486		19	PARENTHESIZED NUMBER NINETEEN
U+2487		20	PARENTHESIZED NUMBER TWENTY
U+2488		1	DIGIT ONE FULL STOP
U+2489		2	DIGIT TWO FULL STOP
U+248A		3	DIGIT THREE FULL STOP
U+248B		4	DIGIT FOUR FULL STOP
U+248C		5	DIGIT FIVE FULL STOP
U+248D		6	DIGIT SIX FULL STOP
U+248E		7	DIGIT SEVEN FULL STOP
U+248F		8	DIGIT EIGHT FULL STOP
U+2490		9	DIGIT NINE FULL STOP
U+2491		10	NUMBER TEN FULL STOP
U+2492		11	NUMBER ELEVEN FULL STOP
U+2493		12	NUMBER TWELVE FULL STOP
U+2494		13	NUMBER THIRTEEN FULL STOP
U+2495		14	NUMBER FOURTEEN FULL STOP
U+2496		15	NUMBER FIFTEEN FULL STOP
U+2497		16	NUMBER SIXTEEN FULL STOP
U+2498		17	NUMBER SEVENTEEN FULL STOP
U+2499		18	NUMBER EIGHTEEN FULL STOP
U+249A		19	NUMBER NINETEEN FULL STOP
U+249B		20	NUMBER TWENTY FULL STOP
U+24EA		0	CIRCLED DIGIT ZERO
U+2776		1	DINGBAT NEGATIVE CIRCLED DIGIT ONE
U+2777		2	DINGBAT NEGATIVE CIRCLED DIGIT TWO
U+2778		3	DINGBAT NEGATIVE CIRCLED DIGIT THREE
U+2779		4	DINGBAT NEGATIVE CIRCLED DIGIT FOUR
U+277A		5	DINGBAT NEGATIVE CIRCLED DIGIT FIVE
U+277B		6	DINGBAT NEGATIVE CIRCLED DIGIT SIX
U+277C		7	DINGBAT NEGATIVE CIRCLED DIGIT SEVEN
U+277D		8	DINGBAT NEGATIVE CIRCLED DIGIT EIGHT

Table 4-6. Numeric Properties (Continued)

Value	Decimal	Number	Name
U+277E		9	DINGBAT NEGATIVE CIRCLED DIGIT NINE
U+277F		10	DINGBAT NEGATIVE CIRCLED NUMBER TEN
U+2780		1	DINGBAT CIRCLED SANS-SERIF DIGIT ONE
U+2781		2	DINGBAT CIRCLED SANS-SERIF DIGIT TWO
U+2782		3	DINGBAT CIRCLED SANS-SERIF DIGIT THREE
U+2783		4	DINGBAT CIRCLED SANS-SERIF DIGIT FOUR
U+2784		5	DINGBAT CIRCLED SANS-SERIF DIGIT FIVE
U+2785		6	DINGBAT CIRCLED SANS-SERIF DIGIT SIX
U+2786		7	DINGBAT CIRCLED SANS-SERIF DIGIT SEVEN
U+2787		8	DINGBAT CIRCLED SANS-SERIF DIGIT EIGHT
U+2788		9	DINGBAT CIRCLED SANS-SERIF DIGIT NINE
U+2789		10	DINGBAT CIRCLED SANS-SERIF NUMBER TEN
U+278A		1	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT ONE
U+278B		2	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT TWO
U+278C		3	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT THREE
U+278D		4	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT FOUR
U+278E		5	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT FIVE
U+278F		6	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT SIX
U+2790		7	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT SEVEN
U+2791		8	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT EIGHT
U+2792		9	DINGBAT NEGATIVE CIRCLED SANS-SERIF DIGIT NINE
U+2793		10	DINGBAT NEGATIVE CIRCLED SANS-SERIF NUMBER TEN
U+3007		0	IDEOGRAPHIC NUMBER ZERO
U+3021		1	HANGZHOU NUMERAL ONE
U+3022		2	HANGZHOU NUMERAL TWO
U+3023		3	HANGZHOU NUMERAL THREE
U+3024		4	HANGZHOU NUMERAL FOUR
U+3025		5	HANGZHOU NUMERAL FIVE
U+3026		6	HANGZHOU NUMERAL SIX
U+3027		7	HANGZHOU NUMERAL SEVEN
U+3028		8	HANGZHOU NUMERAL EIGHT
U+3029		9	HANGZHOU NUMERAL NINE
U+3280		1	CIRCLED IDEOGRAPH ONE
U+3281		2	CIRCLED IDEOGRAPH TWO
U+3282		3	CIRCLED IDEOGRAPH THREE
U+3283		4	CIRCLED IDEOGRAPH FOUR
U+3284		5	CIRCLED IDEOGRAPH FIVE
U+3285		6	CIRCLED IDEOGRAPH SIX
U+3286		7	CIRCLED IDEOGRAPH SEVEN
U+3287		8	CIRCLED IDEOGRAPH EIGHT
U+3288		9	CIRCLED IDEOGRAPH NINE
U+3289		10	CIRCLED IDEOGRAPH TEN
U+FF10	✓	0	FULLWIDTH DIGIT ZERO
U+FF11	✓	1	FULLWIDTH DIGIT ONE
U+FF12	✓	2	FULLWIDTH DIGIT TWO
U+FF13	✓	3	FULLWIDTH DIGIT THREE
U+FF14	✓	4	FULLWIDTH DIGIT FOUR
U+FF15	✓	5	FULLWIDTH DIGIT FIVE
U+FF16	✓	6	FULLWIDTH DIGIT SIX
U+FF17	✓	7	FULLWIDTH DIGIT SEVEN
U+FF18	✓	8	FULLWIDTH DIGIT EIGHT
U+FF19	✓	9	FULLWIDTH DIGIT NINE

4.7 Mirrored

Mirrored is the normative property of character pairs such as parentheses whose images are mirrored horizontally in text that is laid out from right to left. For example, U+0028 LEFT PARENTHESIS is interpreted as *opening parenthesis*; in a left-to-right context, this will appear as “(”, while in a right-to-left context, this will appear as the mirrored glyph “)”. The list of mirrored characters appears in Table 4-7.

Table 4-7. Mirrored Characters

Value	Name
U+0028	LEFT PARENTHESIS
U+0029	RIGHT PARENTHESIS
U+003C	LESS-THAN SIGN
U+003E	GREATER-THAN SIGN
U+005B	LEFT SQUARE BRACKET
U+005D	RIGHT SQUARE BRACKET
U+007B	LEFT CURLY BRACKET
U+007D	RIGHT CURLY BRACKET
U+2045	LEFT SQUARE BRACKET WITH QUILL
U+2046	RIGHT SQUARE BRACKET WITH QUILL
U+207D	SUPERSCRIP ^T LEFT PARENTHESIS
U+207E	SUPERSCRIP ^T RIGHT PARENTHESIS
U+208D	SUBSCRIP ^T LEFT PARENTHESIS
U+208E	SUBSCRIP ^T RIGHT PARENTHESIS
U+2201	COMPLEMENT
U+2202	PARTIAL DIFFERENTIAL
U+2203	THERE EXISTS
U+2204	THERE DOES NOT EXIST
U+2208	ELEMENT OF
U+2209	NOT AN ELEMENT OF
U+220A	SMALL ELEMENT OF
U+220B	CONTAINS AS MEMBER
U+220C	DOES NOT CONTAIN AS MEMBER
U+220D	SMALL CONTAINS AS MEMBER
U+2211	N-ARY SUMMATION
U+2215	DIVISION SLASH
U+2216	SET MINUS
U+221A	SQUARE ROOT
U+221B	CUBE ROOT
U+221C	FOURTH ROOT
U+221D	PROPORTIONAL TO
U+221F	RIGHT ANGLE
U+2220	ANGLE
U+2221	MEASURED ANGLE
U+2222	SPHERICAL ANGLE
U+2224	DOES NOT DIVIDE
U+2226	NOT PARALLEL TO
U+222B	INTEGRAL
U+222C	DOUBLE INTEGRAL
U+222D	TRIPLE INTEGRAL
U+222E	CONTOUR INTEGRAL
U+222F	SURFACE INTEGRAL
U+2230	VOLUME INTEGRAL
U+2231	CLOCKWISE INTEGRAL
U+2232	CLOCKWISE CONTOUR INTEGRAL
U+2233	ANTICLOCKWISE CONTOUR INTEGRAL
U+2239	EXCESS
U+223B	HOMOTHETIC
U+223C	TILDE OPERATOR
U+223D	REVERSED TILDE

Table 4-7. Mirrored Characters (Continued)

Value	Name
U+223E	INVERTED LAZY S
U+223F	SINE WAVE
U+2240	WREATH PRODUCT
U+2241	NOT TILDE
U+2242	MINUS TILDE
U+2243	ASYMPTOTICALLY EQUAL TO
U+2244	NOT ASYMPTOTICALLY EQUAL TO
U+2245	APPROXIMATELY EQUAL TO
U+2246	APPROXIMATELY BUT NOT ACTUALLY EQUAL TO
U+2247	NEITHER APPROXIMATELY NOR ACTUALLY EQUAL TO
U+2248	ALMOST EQUAL TO
U+2249	NOT ALMOST EQUAL TO
U+224A	ALMOST EQUAL OR EQUAL TO
U+224B	TRIPLE TILDE
U+224C	ALL EQUAL TO
U+2252	APPROXIMATELY EQUAL TO OR THE IMAGE OF
U+2253	IMAGE OF OR APPROXIMATELY EQUAL TO
U+2254	COLON EQUALS
U+2255	EQUALS COLON
U+225F	QUESTIONED EQUAL TO
U+2260	NOT EQUAL TO
U+2262	NOT IDENTICAL TO
U+2264	LESS-THAN OR EQUAL TO
U+2265	GREATER-THAN OR EQUAL TO
U+2266	LESS-THAN OVER EQUAL TO
U+2267	GREATER-THAN OVER EQUAL TO
U+2268	LESS-THAN BUT NOT EQUAL TO
U+2269	GREATER-THAN BUT NOT EQUAL TO
U+226A	MUCH LESS-THAN
U+226B	MUCH GREATER-THAN
U+226E	NOT LESS-THAN
U+226F	NOT GREATER-THAN
U+2270	NEITHER LESS-THAN NOR EQUAL TO
U+2271	NEITHER GREATER-THAN NOR EQUAL TO
U+2272	LESS-THAN OR EQUIVALENT TO
U+2273	GREATER-THAN OR EQUIVALENT TO
U+2274	NEITHER LESS-THAN NOR EQUIVALENT TO
U+2275	NEITHER GREATER-THAN NOR EQUIVALENT TO
U+2276	LESS-THAN OR GREATER-THAN
U+2277	GREATER-THAN OR LESS-THAN
U+2278	NEITHER LESS-THAN NOR GREATER-THAN
U+2279	NEITHER GREATER-THAN NOR LESS-THAN
U+227A	PRECEDES
U+227B	SUCCEEDS
U+227C	PRECEDES OR EQUAL TO
U+227D	SUCCEEDS OR EQUAL TO
U+227E	PRECEDES OR EQUIVALENT TO
U+227F	SUCCEEDS OR EQUIVALENT TO
U+2280	DOES NOT PRECEDE
U+2281	DOES NOT SUCCEED
U+2282	SUBSET OF
U+2283	SUPERSET OF
U+2284	NOT A SUBSET OF
U+2285	NOT A SUPERSET OF
U+2286	SUBSET OF OR EQUAL TO
U+2287	SUPERSET OF OR EQUAL TO
U+2288	NEITHER A SUBSET OF NOR EQUAL TO
U+2289	NEITHER A SUPERSET OF NOR EQUAL TO
U+228A	SUBSET OF WITH NOT EQUAL TO
U+228B	SUPERSET OF WITH NOT EQUAL TO
U+228C	MULTISET

Table 4-7. Mirrored Characters (Continued)

Value	Name
U+228F	SQUARE IMAGE OF
U+2290	SQUARE ORIGINAL OF
U+2291	SQUARE IMAGE OF OR EQUAL TO
U+2292	SQUARE ORIGINAL OF OR EQUAL TO
U+2298	CIRCLED DIVISION SLASH
U+22A2	RIGHT TACK
U+22A3	LEFT TACK
U+22A6	ASSERTION
U+22A7	MODELS
U+22A8	TRUE
U+22A9	FORCES
U+22AA	TRIPLE VERTICAL BAR RIGHT TURNSTILE
U+22AB	DOUBLE VERTICAL BAR DOUBLE RIGHT TURNSTILE
U+22AC	DOES NOT PROVE
U+22AD	NOT TRUE
U+22AE	DOES NOT FORCE
U+22AF	NEGATED DOUBLE VERTICAL BAR DOUBLE RIGHT TURNSTILE
U+22B0	PRECEDES UNDER RELATION
U+22B1	SUCCEEDS UNDER RELATION
U+22B2	NORMAL SUBGROUP OF
U+22B3	CONTAINS AS NORMAL SUBGROUP
U+22B4	NORMAL SUBGROUP OF OR EQUAL TO
U+22B5	CONTAINS AS NORMAL SUBGROUP OR EQUAL TO
U+22B6	ORIGINAL OF
U+22B7	IMAGE OF
U+22B8	MULTIMAP
U+22BE	RIGHT ANGLE WITH ARC
U+22BF	RIGHT TRIANGLE
U+22C9	LEFT NORMAL FACTOR SEMIDIRECT PRODUCT
U+22CA	RIGHT NORMAL FACTOR SEMIDIRECT PRODUCT
U+22CB	LEFT SEMIDIRECT PRODUCT
U+22CC	RIGHT SEMIDIRECT PRODUCT
U+22CD	REVERSED TILDE EQUALS
U+22D0	DOUBLE SUBSET
U+22D1	DOUBLE SUPERSET
U+22D6	LESS-THAN WITH DOT
U+22D7	GREATER-THAN WITH DOT
U+22D8	VERY MUCH LESS-THAN
U+22D9	VERY MUCH GREATER-THAN
U+22DA	LESS-THAN EQUAL TO OR GREATER-THAN
U+22DB	GREATER-THAN EQUAL TO OR LESS-THAN
U+22DC	EQUAL TO OR LESS-THAN
U+22DD	EQUAL TO OR GREATER-THAN
U+22DE	EQUAL TO OR PRECEDES
U+22DF	EQUAL TO OR SUCCEEDS
U+22E0	DOES NOT PRECEDE OR EQUAL
U+22E1	DOES NOT SUCCEED OR EQUAL
U+22E2	NOT SQUARE IMAGE OF OR EQUAL TO
U+22E3	NOT SQUARE ORIGINAL OF OR EQUAL TO
U+22E4	SQUARE IMAGE OF OR NOT EQUAL TO
U+22E5	SQUARE ORIGINAL OF OR NOT EQUAL TO
U+22E6	LESS-THAN BUT NOT EQUIVALENT TO
U+22E7	GREATER-THAN BUT NOT EQUIVALENT TO
U+22E8	PRECEDES BUT NOT EQUIVALENT TO
U+22E9	SUCCEEDS BUT NOT EQUIVALENT TO
U+22EA	NOT NORMAL SUBGROUP OF
U+22EB	DOES NOT CONTAIN AS NORMAL SUBGROUP
U+22EC	NOT NORMAL SUBGROUP OF OR EQUAL TO
U+22ED	DOES NOT CONTAIN AS NORMAL SUBGROUP OR EQUAL
U+22F0	UP RIGHT DIAGONAL ELLIPSIS
U+22F1	DOWN RIGHT DIAGONAL ELLIPSIS

Table 4-7. Mirrored Characters (Continued)

Value	Name
U+2308	LEFT CEILING
U+2309	RIGHT CEILING
U+230A	LEFT FLOOR
U+230B	RIGHT FLOOR
U+2320	TOP HALF INTEGRAL
U+2321	BOTTOM HALF INTEGRAL
U+2329	LEFT-POINTING ANGLE BRACKET
U+232A	RIGHT-POINTING ANGLE BRACKET
U+3008	LEFT ANGLE BRACKET
U+3009	RIGHT ANGLE BRACKET
U+300A	LEFT DOUBLE ANGLE BRACKET
U+300B	RIGHT DOUBLE ANGLE BRACKET
U+300C	LEFT CORNER BRACKET
U+300D	RIGHT CORNER BRACKET
U+300E	LEFT WHITE CORNER BRACKET
U+300F	RIGHT WHITE CORNER BRACKET
U+3010	LEFT BLACK LENTICULAR BRACKET
U+3011	RIGHT BLACK LENTICULAR BRACKET
U+3014	LEFT TORTOISE SHELL BRACKET
U+3015	RIGHT TORTOISE SHELL BRACKET
U+3016	LEFT WHITE LENTICULAR BRACKET
U+3017	RIGHT WHITE LENTICULAR BRACKET
U+3018	LEFT WHITE TORTOISE SHELL BRACKET
U+3019	RIGHT WHITE TORTOISE SHELL BRACKET
U+301A	LEFT WHITE SQUARE BRACKET
U+301B	RIGHT WHITE SQUARE BRACKET

4.8 Unicode 1.0 Names

The *Unicode 1.0 character name* is an informative property of the characters defined in the Unicode Standard Version 1.0. The names of Unicode characters were changed in the process of merging with ISO/IEC 10646. The 1.0 character names can be obtained from the CD-ROM accompanying the standard or from the ftp site. (See also *Appendix D, Cumulative Changes*.)

4.9 Mathematical Property

The mathematical property is an informative property of characters that are used as operators in mathematical formulas. The mathematical property may be useful in algorithms that deal with the display of mathematical text and formulas. However, a number of these characters have multiple usages and may also occur with non-mathematical semantics. For example, U+002D HYPHEN-MINUS may also be used as a hyphen—and not as a mathematical minus sign. Other characters, including some alphabetic, numeric, punctuation, spaces, arrows, and geometric shapes, are also used in mathematical expressions, but are even more dependent on the context for their identification. The following list of Unicode characters have the mathematical property.

0028..002B, 002D, 002F, 003C..003E, 005B..005E, 007B..007E, 00AC, 00B1,
 00D7, 00F7, 20A6, 2032..2034, 2044, 20D2..20E1, 2190..2194, 21D2, 21D4,
 2200..22F1, 2308..230B, 2320..2321, 2329..232A