

# UnicodeData Template

B/unict

## *Purpose*

To tabulate casings in the character set.

---

B/unict.§1 Source; §2 ZSCII Casing Tables; §3 Small Unicode Casing Tables; §4 Large Unicode Casing Tables

---

§1. **Source.** When this section is included, exactly one of the following constants is defined:

- (a) **ZSCII\_Tables**, meaning that we will use ZSCII as the character set for characters in indexed text strings.
- (b) **Small\_Unicode\_Tables**, meaning that we will use Unicode but store only single-byte characters, so that only the codes 0 to 255 are valid: in effect ISO Latin-1.
- (c) **Large\_Unicode\_Tables**, meaning that we will use Unicode and store two-byte characters, so that all of Unicode in the range 0 to 65535 are valid.

Whichever is defined, we must create two arrays:

- (i) **CharCasingChart0**, a table indicating lower-case letters with transitions to convert them to upper case;
- (ii) **CharCasingChart1**, vice versa.

Each array is a sequence of three-word records, consisting of the start of a character range, the size of the range (the number of characters in it), and the numerical offset to convert to the opposite case. For instance, the sequence (97,26,-32) means the 26 lower-case letters “a” to “z”, and marks them as convertible to upper case by subtracting 32 from the character code (so “a”, 97, becomes “A”, 65). If the size of the range is negative, this indicates that only every alternate code is valid. (This makes for efficient storage since there are large parts of the Unicode number-space in which upper and lower case letters alternate.)

An offset of UNIC\_NCT means no case change is possible; and any character not included in the ranges below is not a letter.

```
Constant UNIC_NCT = 10000; ! Safe as highest case-change delta is 8383
```

## §2. ZSCII Casing Tables.

```
#IFDEF ZSCII_Tables;
Array CharCasingChart0 -->
    $0061 ( 26) (   -32) $009b (  3) (      3) $00a1 (  1) (UNIC_NCT)
    $00a4 (  2) (      3) $00a6 (  1) (UNIC_NCT) $00a9 (  6) (      6)
    $00b5 (  5) (      5) $00bf (  5) (      5) $00c9 ( -3) (      1)
    $00cd (  3) (      3) $00d3 ( -3) (      1) $00d7 (  2) (      2)
    $00dc (  1) (      1) $0000
;
Array CharCasingChart1 -->
    $0041 ( 26) (    32) $009e (  3) (     -3) $00a7 (  2) (     -3)
    $00af (  6) (     -6) $00ba (  5) (     -5) $00c4 (  5) (     -5)
    $00ca ( -3) (     -1) $00d0 (  3) (     -3) $00d4 ( -3) (     -1)
    $00d9 (  2) (     -2) $00dd (  1) (     -1) $0000
;
#ENDIF; ! ZSCII_Tables
```

## §3. Small Unicode Casing Tables.

```
#IFDEF Small_Unicode_Tables;
Array CharCasingChart0 -->
    $0061 ( 26) (   -32) $00aa (  1) (UNIC_NCT) $00b5 (  1) (UNIC_NCT) $00ba (  1) (UNIC_NCT)
    $00df (  1) (UNIC_NCT) $00e0 ( 23) (   -32) $00f8 (  7) (   -32) $00ff (  1) (UNIC_NCT)
    $0000
;
Array CharCasingChart1 -->
    $0041 ( 26) (   32) $00c0 ( 23) (   32) $00d8 (  7) (   32) $0000
;
#ENDIF; ! Small_Unicode_Tables
```

## §4. Large Unicode Casing Tables.

```
#IFDEF Large_Unicode_Tables;
Array CharCasingChart0 -->
    $0061 ( 26) (   -32) $00aa (  1) (UNIC_NCT) $00b5 (  1) (   743) $00ba (  1) (UNIC_NCT)
    $00df (  1) (UNIC_NCT) $00e0 ( 23) (   -32) $00f8 (  7) (   -32) $00ff (  1) (   121)
    $0101 ( -47) (   -1) $0131 (  1) (  -232) $0133 ( -5) (   -1) $0138 (  1) (UNIC_NCT)
    $013a ( -15) (   -1) $0149 (  1) (UNIC_NCT) $014b ( -45) (   -1) $017a ( -5) (   -1)
    $017f (  1) (  -300) $0180 (  1) (UNIC_NCT) $0183 ( -3) (   -1) $0188 (  1) (   -1)
    $018c (  1) (   -1) $018d (  1) (UNIC_NCT) $0192 (  1) (   -1) $0195 (  1) (   97)
    $0199 (  1) (   -1) $019a (  2) (UNIC_NCT) $019e (  1) (   130) $01a1 ( -5) (   -1)
    $01a8 (  1) (   -1) $01aa (  2) (UNIC_NCT) $01ad (  1) (   -1) $01b0 (  1) (   -1)
    $01b4 ( -3) (   -1) $01b9 (  1) (   -1) $01ba (  1) (UNIC_NCT) $01bd (  1) (   -1)
    $01be (  1) (UNIC_NCT) $01bf (  1) (   56) $01c6 (  1) (   -2) $01c9 (  1) (   -2)
    $01cc (  1) (   -2) $01ce ( -15) (   -1) $01dd (  1) (  -79) $01df ( -17) (   -1)
    $01f0 (  1) (UNIC_NCT) $01f3 (  1) (   -2) $01f5 (  1) (   -1) $01f9 ( -39) (   -1)
    $0221 (  1) (UNIC_NCT) $0223 ( -17) (   -1) $0234 (  3) (UNIC_NCT) $0250 (  3) (UNIC_NCT)
    $0253 (  1) (  -210) $0254 (  1) (  -206) $0255 (  1) (UNIC_NCT) $0256 (  2) (  -205)
    $0258 (  1) (UNIC_NCT) $0259 (  1) (  -202) $025a (  1) (UNIC_NCT) $025b (  1) (  -203)
    $025c (  4) (UNIC_NCT) $0260 (  1) (  -205) $0261 (  2) (UNIC_NCT) $0263 (  1) (  -207)
    $0264 (  4) (UNIC_NCT) $0268 (  1) (  -209) $0269 (  1) (  -211) $026a (  5) (UNIC_NCT)
    $026f (  1) (  -211) $0270 (  2) (UNIC_NCT) $0272 (  1) (  -213) $0273 (  2) (UNIC_NCT)
    $0275 (  1) (  -214) $0276 ( 10) (UNIC_NCT) $0280 (  1) (  -218) $0281 (  2) (UNIC_NCT)
    $0283 (  1) (  -218) $0284 (  4) (UNIC_NCT) $0288 (  1) (  -218) $0289 (  1) (UNIC_NCT)
    $028a (  2) (  -217) $028c (  6) (UNIC_NCT) $0292 (  1) (  -219) $0293 ( 29) (UNIC_NCT)
    $0390 (  1) (UNIC_NCT) $03ac (  1) (   -38) $03ad (  3) (   -37) $03b0 (  1) (UNIC_NCT)
    $03b1 ( 17) (   -32) $03c2 (  1) (   -31) $03c3 (  9) (   -32) $03cc (  1) (   -64)
    $03cd (  2) (   -63) $03d0 (  1) (   -62) $03d1 (  1) (   -57) $03d5 (  1) (   -47)
    $03d6 (  1) (   -54) $03d7 (  1) (UNIC_NCT) $03d9 ( -23) (   -1) $03f0 (  1) (   -86)
    $03f1 (  1) (   -80) $03f2 (  1) (    7) $03f3 (  1) (UNIC_NCT) $03f5 (  1) (   -96)
    $03f8 (  1) (   -1) $03fb (  1) (   -1) $0430 ( 32) (   -32) $0450 ( 16) (   -80)
    $0461 ( -33) (   -1) $048b ( -53) (   -1) $04c2 ( -13) (   -1) $04d1 ( -37) (   -1)
    $04f9 (  1) (   -1) $0501 ( -15) (   -1) $0561 ( 38) (  -48) $0587 (  1) (UNIC_NCT)
    $1d00 ( 44) (UNIC_NCT) $1d62 ( 10) (UNIC_NCT) $1e01 (-149) (   -1) $1e96 (  5) (UNIC_NCT)
    $1e9b (  1) (  -59) $1ea1 ( -89) (   -1) $1f00 (  8) (    8) $1f10 (  6) (    8)
    $1f20 (  8) (    8) $1f30 (  8) (    8) $1f40 (  6) (    8) $1f50 (  1) (UNIC_NCT)
    $1f51 (  1) (    8) $1f52 (  1) (UNIC_NCT) $1f53 (  1) (    8) $1f54 (  1) (UNIC_NCT)
    $1f55 (  1) (    8) $1f56 (  1) (UNIC_NCT) $1f57 (  1) (    8) $1f60 (  8) (    8)
    $1f70 (  2) (   74) $1f72 (  4) (   86) $1f76 (  2) (  100) $1f78 (  2) (  128)
```

```

$1f7a ( 2) ( 112) $1f7c ( 2) ( 126) $1f80 ( 8) ( 8) $1f90 ( 8) ( 8)
$1fa0 ( 8) ( 8) $1fb0 ( 2) ( 8) $1fb2 ( 1) (UNIC_NCT) $1fb3 ( 1) ( 9)
$1fb4 ( -3) (UNIC_NCT) $1fb7 ( 1) (UNIC_NCT) $1fbe ( 1) ( -7205) $1fc2 ( 1) (UNIC_NCT)
$1fc3 ( 1) ( 9) $1fc4 ( -3) (UNIC_NCT) $1fc7 ( 1) (UNIC_NCT) $1fd0 ( 2) ( 8)
$1fd2 ( 2) (UNIC_NCT) $1fd6 ( 2) (UNIC_NCT) $1fe0 ( 2) ( 8) $1fe2 ( 3) (UNIC_NCT)
$1fe5 ( 1) ( 7) $1fe6 ( 2) (UNIC_NCT) $1ff2 ( 1) (UNIC_NCT) $1ff3 ( 1) ( 9)
$1ff4 ( -3) (UNIC_NCT) $1ff7 ( 1) (UNIC_NCT) $2071 ( 1) (UNIC_NCT) $207f ( 1) (UNIC_NCT)
$210a ( 1) (UNIC_NCT) $210e ( 2) (UNIC_NCT) $2113 ( 1) (UNIC_NCT) $212f ( 1) (UNIC_NCT)
$2134 ( 1) (UNIC_NCT) $2139 ( 1) (UNIC_NCT) $213d ( 1) (UNIC_NCT) $2146 ( 4) (UNIC_NCT)
$fb00 ( 7) (UNIC_NCT) $fb13 ( 5) (UNIC_NCT) $ff41 ( 26) ( -32) $0000

```

```
;
```

```
Array CharCasingChart1 -->
```

```

$0041 ( 26) ( 32) $00c0 ( 23) ( 32) $00d8 ( 7) ( 32) $0100 ( -47) ( 1)
$0130 ( 1) ( -199) $0132 ( -5) ( 1) $0139 ( -15) ( 1) $014a ( -45) ( 1)
$0178 ( 1) ( -121) $0179 ( -5) ( 1) $0181 ( 1) ( 210) $0182 ( -3) ( 1)
$0186 ( 1) ( 206) $0187 ( 1) ( 1) $0189 ( 2) ( 205) $018b ( 1) ( 1)
$018e ( 1) ( 79) $018f ( 1) ( 202) $0190 ( 1) ( 203) $0191 ( 1) ( 1)
$0193 ( 1) ( 205) $0194 ( 1) ( 207) $0196 ( 1) ( 211) $0197 ( 1) ( 209)
$0198 ( 1) ( 1) $019c ( 1) ( 211) $019d ( 1) ( 213) $019f ( 1) ( 214)
$01a0 ( -5) ( 1) $01a6 ( 1) ( 218) $01a7 ( 1) ( 1) $01a9 ( 1) ( 218)
$01ac ( 1) ( 1) $01ae ( 1) ( 218) $01af ( 1) ( 1) $01b1 ( 2) ( 217)
$01b3 ( -3) ( 1) $01b7 ( 1) ( 219) $01b8 ( 1) ( 1) $01bc ( 1) ( 1)
$01c4 ( 1) ( 2) $01c7 ( 1) ( 2) $01ca ( 1) ( 2) $01cd ( -15) ( 1)
$01de ( -17) ( 1) $01f1 ( 1) ( 2) $01f4 ( 1) ( 1) $01f6 ( 1) ( -97)
$01f7 ( 1) ( -56) $01f8 ( -39) ( 1) $0220 ( 1) ( -130) $0222 ( -17) ( 1)
$0386 ( 1) ( 38) $0388 ( 3) ( 37) $038c ( 1) ( 64) $038e ( 2) ( 63)
$0391 ( 17) ( 32) $03a3 ( 9) ( 32) $03d2 ( 3) (UNIC_NCT) $03d8 ( -23) ( 1)
$03f4 ( 1) ( -60) $03f7 ( 1) ( 1) $03f9 ( 1) ( -7) $03fa ( 1) ( 1)
$0400 ( 16) ( 80) $0410 ( 32) ( 32) $0460 ( -33) ( 1) $048a ( -53) ( 1)
$04c0 ( 1) (UNIC_NCT) $04c1 ( -13) ( 1) $04d0 ( -37) ( 1) $04f8 ( 1) ( 1)
$0500 ( -15) ( 1) $0531 ( 38) ( 48) $10a0 ( 38) (UNIC_NCT) $1e00 ( -149) ( 1)
$1ea0 ( -89) ( 1) $1f08 ( 8) ( -8) $1f18 ( 6) ( -8) $1f28 ( 8) ( -8)
$1f38 ( 8) ( -8) $1f48 ( 6) ( -8) $1f59 ( -7) ( -8) $1f68 ( 8) ( -8)
$1fb8 ( 2) ( -8) $1fba ( 2) ( -74) $1fc8 ( 4) ( -86) $1fd8 ( 2) ( -8)
$1fda ( 2) ( -100) $1fe8 ( 2) ( -8) $1fea ( 2) ( -112) $1fec ( 1) ( -7)
$1ff8 ( 2) ( -128) $1ffa ( 2) ( -126) $2102 ( 1) (UNIC_NCT) $2107 ( 1) (UNIC_NCT)
$210b ( 3) (UNIC_NCT) $2110 ( 3) (UNIC_NCT) $2115 ( 1) (UNIC_NCT) $2119 ( 5) (UNIC_NCT)
$2124 ( 1) (UNIC_NCT) $2126 ( 1) ( -7517) $2128 ( 1) (UNIC_NCT) $212a ( 1) ( -8383)
$212b ( 1) ( -8262) $212c ( 2) (UNIC_NCT) $2130 ( 2) (UNIC_NCT) $2133 ( 1) (UNIC_NCT)
$213e ( 2) (UNIC_NCT) $2145 ( 1) (UNIC_NCT) $ff21 ( 26) ( 32) $0000

```

```
;
```

```
#ENDIF; ! Large_Unicode_Tables
```