Greek and hyperref

Günter Milde

2020/10/30

On 2010-11-05, Heiko Oberdiek wrote in comp.text.tex:

> \pdfstringdef (converting TeX code to PDF strings for bookmarks) supports NFSS2 and needs active characters. Encoding stuff based on the internal font machinery of TeX (letters with catcode 11 or 12, ligatures) does not work, because the strings don't reach TeX's stomach.

The *greek-fontenc* package allows input of Greek characters in a way that "reaches TeX's stomach" and hence works in both, the main document as well as in PDF strings (e.g. bookmarks). Hyperref's "puenc.def" font encoding file defines LICR macros for monotonic Greek (Greek characters of the "Greek and Coptic" unicode block).

All utf8-encoded literal Unicode characters work in PDF strings. With *greek-fontenc* and *greek-inputenc*, this enables use of all Greek character in text and PDF strings.

1 Transcription: λογος, λογος

\textgreek + LGR transcription or Greek language (babel) + LGR transcription: In the PDF-bookmark are Latin letters instead of Greek ones.

2 Macros: λογος, λογος, λογος

textalpha package with \textalpha ... \textomega macros or *alphabeta* package with \alpha ... \Omega macros. With 8-bit TeX engines (pdflatex), literal Greek Unicode characters are converted to LICR Macros, too.

Works, if the `unicode` or `pdfencoding=auto` option is given to *hyperref*.[1]

The generic short macros from the *alphabeta* package result in hyperref warnings. See alphabeta-doc.pdf and alphabeta-doc.tex for details and workarounds.

[1] With the "xpdf" viewer, Greek letters are not shown in PDF bookmarks.

3 LGR + Macros: λογος

LICR-macro input works also if the font encoding is LGR.

4 Kerning: AϒΛ AϒΛ AϒΛ

Kerning is impossible if the font encoding is switched for every single character. To fix this, wrap the Greek part in a command switching to a font encoding supporting Greek, either `\ensuregreek{...}` (with package *textalpha* or *Babel*) or `\foreignlanguage{greek}{...}` (with *Babel*).

5 Literal Unicode input

The following subsection headings contain all characters from the "Greek and Coptic" and "Greek Extended" Unicode Blocks that are supported by the LGR font encoding.

5.1 ʹ„; ΄ ΅ ΆΑ·ΈΕΉΙΌΟΎΥΏΰἀΑΒΓΔΕΖΗΘΙΚΛΜΝ

5.2 ΞΟΠΡΣΤΥΦΧΨΩΪΫῨῩϜFⅎ

Greek and Coptic Unicode block: punctuation and uppercase letters

5.3 άέήίύαβγδεζηθικλμν

5.4 ξοπρςστυφχψωϊϋόύώβθφπῩῨϝϜɬↄ

Greek and Coptic Unicode block: lowercase letters

5.5 ἀἁἂἃἄἅἆἇ᾽Α῾Α῾Α῾Α῾Α῾Α῾Α῾Α

5.6 ἐἑἒἓἔἕ᾽Ε῾Ε῾Ε῾Ε῾Ε῾Ε

5.7 ἠἡἢἣἤἥἦἧ᾽Η῾Η῾Η῾Η῾Η῾Η῾Η῾Η

5.8 ἰἱἲἳἴἵἶἷ᾽Ι῾Ι῾Ι῾Ι῾Ι῾Ι῾Ι῾Ι

5.9 ὀὁὂὃὄὅ᾽Ο῾Ο῾Ο῾Ο῾Ο῾Ο

5.10 ὐὑὒὓὔὕὖὗ῾Υ῾Υ῾Υ῾Υ

5.11 ὠὡὢὣὤὥὦὧ᾽Ω῾Ω῾Ω῾Ω῾Ω῾Ω῾Ω῾Ω

5.12 ὰάὲέὴήὶίὸόὺύὼώ

5.13 ᾀᾁᾂᾃᾄᾅᾆᾇ᾽Α῾Α῾Α῾Α῾Α῾Α῾Α῾Α

5.14 ᾐᾑᾒᾓᾔᾕᾖᾗ᾽Η῾Η῾Η῾Η῾Η῾Η῾Η῾Η

5.15 ᾠᾡᾢᾣᾤᾥᾦᾧ᾽Ω῾Ω῾Ω῾Ω῾Ω῾Ω῾Ω῾Ω

5.16 ᾰᾱᾲᾳᾴᾶᾷᾸᾹᾺΆΑ῾᾽῾

5.17 ῀῁ῂῃῄῆῇῈΈ῾Η῾ΗΗ῾῍῎

5.18 ῒΐῖῗῘΙ῾Ῑ῾Ι῾῏῞

5.19 ῠῡῢΰῤῥῦῧῨΥ῾Ῡ῾Υ῾Υ῾Ρ῝῟

5.20 ῲῳῴῶῷῸ῾Ο῾Ω῾ΩΩ῾῭΅

Greek Extended Unicode block: Input as literal precomposed Unicode character works fine.

6 LICR command input

6.1 ´¸ ;´ ¨ ΄Α·Ε΄Η΄Ι΄Ο΄Υ΄Ω΄ΑΒΓΔΕΖΗΘΙΚΛΜΝ

6.2 ΞΟΠΡΣΤΥΦΧΨΩΪΫ

6.3 άέήίΰαβγδεζηθικλμν

6.4 ξοπρςστυφχψωϊϋόύώβϑφπϚϛϜϝFϝϟϟℒℜ

Currently, there is no hyperref support for LICR input with non-standard accents or combined diacritics characters.

6.5 ἀἁἂἃἄἅἆἇᾈᾉᾊᾋᾌᾍᾎᾏ

6.6 ἐἑἒἓἔἕ

6.7 ἠἡἢἣἤἥἦἧᾘᾙᾚᾛᾜᾝᾞᾟ

6.8 ἰἱἲἳἴἵἶἷ

6.9 ὀὁὂὃὄὅ

6.10 ὐὑὒὓὔὕὖὗ

6.11 ὠὡὢὣὤὥὦὧᾨᾩᾪᾫᾬᾭᾮᾯ

6.12 ὰάὲέὴήὶίὸόὺύὼώ

6.13 ᾀᾁᾂᾃᾄᾅᾆᾇᾰᾱᾲᾳᾴᾶᾷ

6.14 ᾐᾑᾒᾓᾔᾕᾖᾗῂῃῄῆῇ

6.15 ᾠᾡᾢᾣᾤᾥᾦᾧῲῳῴῶῷ

6.16 ᾸᾹᾺΆᾼ

6.17 ῈΈῊΉῌ

6.18 ῘῙῚΊ

6.19 ῨῩῪΎῬ

6.20 ῸΌῺΏῼ

7 Makeuppercase

According to Greek typesetting conventions, diacritics (except the dialytika) are dropped in ALL CAPS.

However, `\Makeuppercase` is not supported in PDF-strings, so we do not need to care for this.

8 Conclusion

For Greek text parts in section headers use either literal Unicode characters[2] or macros. For proper kerning in the main document, combine this with the `\textgreek` or `\foreignlanguage{greek}` macros.

[2] Pre-composed Unicode characters must be used for multi-accented letters or non-standard accents. Combining Unicode characters do not work with inputenc and 8-bit LaTeX – this is a general restriction.