Greek and hyperref

Günter Milde

2020/10/30

Contents

1

6 Makeuppercase 5

7 Conclusion 6

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.

With the *textalpha* package, Greek letters can be used without explicit change of the font encoding or Babel language (with some caveats, see textalpha-doc.pdf). For correct hyphenation and other fixes, mark Greek text parts with the Babel language `greek`. There should be no space around a language switch: |λογος|.

1 Transcription: λογος, λογος

Text input using the Latin transliteration: In the PDF-bookmark are Latin letters instead of Greek ones.

2   Macros: λογος, λογος, λογος, λογος,

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

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

3   Kerning: ΑΥΛ ΑΥΛ ΑΥΛ

Kerning is impeded 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*).

4   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.

4.1   ΄.,; ΄ ̈ ΆꞏΈΉΪΟΎꞌΩϊΑΒΓΔΕΖΗΘΙΚΛΜΝ

4.2   ΞΟΠΡΣΤΥΦΧΨΩΪῪϘϹϜϞ

Greek and Coptic Unicode block: punctuation and uppercase letters

4.3   άέήίϋαβγδεζηϑικλμν

4.4   ξοπρϛστυφχψωϊϋόύώϐβϑφπϙϟϝϟϡϰ

Greek and Coptic Unicode block: lowercase letters

---

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

4.5 ἀάᾶᾶᾶᾶᾶ᾽Α᾽Α᾽Α᾽Α᾽Α᾽Α᾽Α᾽Α

4.6 ἐέὲὲὲὲ᾽Ε᾽Ε᾽Ε᾽Ε᾽Ε᾽Ε

4.7 ἠήῆῆῆῆῆ᾽Η᾽Η᾽Η᾽Η᾽Η᾽Η᾽Η᾽Η

4.8 ἰΐῖῖῖῖ᾽Ι᾽Ι᾽Ι᾽Ι᾽Ι᾽Ι᾽Ι

4.9 ὀόὄὄὄὄ᾽Ο᾽Ο᾽Ο᾽Ο᾽Ο᾽Ο

4.10 ὐύὒὒὒὒὒ᾽Υ᾽Υ᾽Υ᾽Υ

4.11 ὠώὢὢὢὢῶῶ᾽Ω᾽Ω᾽Ω᾽Ω᾽Ω᾽Ω᾽Ω᾽Ω

4.12 ἀάὲέῆήὶίὸόὺύὼώ

4.13 ἀᾀᾰᾰᾰᾰᾰᾰ᾽Αι᾽Αι᾽Αι᾽Αι᾽Αι᾽Αι᾽Αι᾽Αι

4.14 ἠήῆῆῆῆῆῆ᾽Ηι᾽Ηι᾽Ηι᾽Ηι᾽Ηι᾽Ηι᾽Ηι᾽Ηι

4.15 ὠώῶῶῶῶῶῶ᾽Ωι᾽Ωι᾽Ωι᾽Ωι᾽Ωι᾽Ωι᾽Ωι᾽Ωι

4.16 ᾰᾱᾰᾰᾰᾰᾰᾸᾹ᾽Α᾽ΑΑι᾽ι᾽

4.17 ῀῾ἠηήῆῆ᾽Ε᾽Ε᾽Η᾽ΗΗι῀῾῀

4.18 ῑῑῑῑῐῙῑῙΙ᾽Ι῀῾῀

4.19 ῠῡὺῠῤῥῦῦΫῩ᾽Υ᾽ΥΡ῾῾῾

4.20 ὠώῴῶῶ᾽Ο᾽Ο᾽Ω᾽ΩΩι῾῾

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

5 LICR command input

5.1 ´¸;᾽ ᾿ Ά·ΈΉΊΌΎΏᾳΑΒΓΔΕΖΗΘΙΚΛΜΝ

5.2 ΞΟΠΡΣΤΥΦΧΨΩΪΫ

5.3 άέήίῢαβγδεζηθικλμν

5.4 ξοπρςστυφχψωϊϋόύώβϑφπϰϱϖϝϛＦϜϟϟϰϰ

textalpha load definitions for LICR input with non-standard accents or combined diacritics characters also for PU (hyperref).

5.5 ἀάᾶᾶᾶᾶᾶᾶ᾿Α᾿Α῍Α῍Α῎Α῎Α῏Α῏Α

5.6 ἐέἒἓἔἕ᾿Ε῍Ε῍Ε῎Ε῎Ε

5.7 ἠήἢἣἤἥἦἧ᾿Η῍Η῍Η῎Η῎Η῏Η῏Η

5.8 ἰίἲἳἴἵἶἷ᾿Ι῍Ι῍Ι῎Ι῎Ι῏Ι῏Ι

5.9 ὀόὂὃὄὅ᾿Ο῍Ο῍Ο῎Ο῎Ο

5.10 ὐὑὒὓὔὕὖὗ῾Υ῍Υ῎Υ῏Υ

5.11 ὠώὢὣὤὥὦὧ᾿Ω῾Ω῍Ω῍Ω῎Ω῎Ω῏Ω

5.12 ὰάὲέὴήὶίὸόὺύὼώ

5.13 ᾀᾁᾂᾃᾄᾅᾆᾇ᾿Αͺ῾Αͺ῍Αͺ῍Αͺ῎Αͺ῎Αͺ῏Αͺ῏Αͺ

5.14 ᾐᾑᾒᾓᾔᾕᾖᾗ᾿Ηͺ῾Ηͺ῍Ηͺ῍Ηͺ῎Ηͺ῎Ηͺ῏Ηͺ῏Ηͺ

5.15 ᾠᾡᾢᾣᾤᾥᾦᾧ᾿Ωͺ῾Ωͺ῍Ωͺ῍Ωͺ῎Ωͺ῎Ωͺ῏Ωͺ῏Ωͺ

5.16 ᾰᾱᾲᾳᾴᾶᾷᾸᾹᾺΆᾼͺ᾿᾿ͺ

5.17 ῀ῂῃῄῆῇῈΈῊΉῌͺ῍῎῏

5.18 ῐῑῒΐῖῗῘῙῚΊ῝῞῟

5.19 ῠῡῢΰῤῥῦῧῨῩῪΎῬ῭΅`

5.20 ῲῳῴῶῷῸΌῺΏῼͺ΅

6 Makeuppercase

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

The LaTeX \Makeuppercase implementation changed fundamentally with the release in 06/2022.[2] \Makeuppercase is now supported PDF-strings. Greek typesetting rules are ony applied if the text language is set to greek with Babel's \setlanguage or \foreignlanguage.

The changes broke the support for upcasing of polutonic Greek (combined accents) with LICR input and of Greek with the LGR Latin transliteration. Restoring support is currently (2023/02/10) worked on.

6.1 AAAAAAAAAAAAAAA

Makeuppercase is not supported in PDF strings with LaTeX versions older than 06/2022.

---

[2]This change cannot be reverted with the rollback mechanism "for technical reasons".

7 Conclusion

For Greek text parts in section headers use either literal Unicode characters[3] or macros. For proper kerning und upcasing in the main document, set the language with \foreignlanguage{greek}.

---

[3]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.